SCANNED, # 8

12 FEB 2002

FORM PTO-1390 (Modified) (REV 5-93) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: 065691-0270

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.5)
PCT FR00/02319	08/11/2000	Unassigned 10/049372

PRIORITY DATE CLAIMED: 08/12/1999

TITLE OF INVENTION: Odorant-Binding Human Proteins Fixing Hydrophobic Ligands: Polypeptides and Polynucleotides Coding for said polypeptides and Uses Thereof

APPLICANT(S) FOR DO/EO/US: Gilles PITIOT, Eric LACAZETTE and Francoise GACHON

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   - ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   - ☒ has been transmitted by the International Bureau.
   - ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   - ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   - ☐ have been transmitted by the International Bureau.
   - ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   - ☒ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).
11. ☐ Applicant claims small entity status under 37 CFR 1.27 .

Items 12. to 17. below concern other document(s) or information included:

12. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
13. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
14. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
15. ☐ A substitute specification.
16. ☐ A change of power of attorney and/or address letter.
17. ☒ Other items or information: Application Data Sheet, Paper Copy of Sequence Listing, and French Language Patent Application Specification

002.708020.1 FORM PTO-1390 (Modified)

SCANNED, #8

JC13 Rec'd PCT/PTO 12 FEB 2002

U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.50)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
Unassigned 10/049372	PCT/FR00/02319	065691-0270

18. ☒ The following fees are submitted:	CALCULATIONS	PTO USE ONLY
Basic National Fee (37 CFR 1.492(a)(1)-(5): Search Report has been prepared by the EPO or JPO.....$890.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482).....$710.00		
No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2) $740.00		
Neither international preliminary examination fee (37 CFR 1.482) nor International search fee (37 CFR 1.445(a)(2)) paid to USPTO $1,040.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) $100.00		
ENTER APPROPRIATE BASIC FEE AMOUNT =	$890.00	
Surcharge of $130.00 for furnishing the oath or declaration later than 20 Months from the earliest claimed priority date (37 CFR 1.492(e))		

Claims	Number Filed		Included in Basic Fee		Extra Claims		Rate		
Total Claims	51	-	20	=	31	×	$18.00	$558.00	
Independent Claims	5	-	3	=	2	×	$84.00	$168.00	
Multiple dependent claim(s) (if applicable)							$280.00	$280.00	
TOTAL OF ABOVE CALCULATIONS =								$1896.00	
Reduction by ½ for filing by small entity, if applicable.								$0.00	
SUBTOTAL =								$1896.00	
Processing fee of $130.00 for furnishing English translation later the 20 months from the earliest claimed priority date (37 CFR 1.492(f). +									
TOTAL NATIONAL FEE =								$1896.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). $40.00 per property +									
TOTAL FEES ENCLOSED =								$1896.00	

				Amount to be:	
				refunded	$
				charged	$

a. ☒ A check in the amount of $1896.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. 19-0741 in the amount of $0.00 to the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 19-0741. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Foley & Lardner
Customer Number: 22428

22428

PATENT TRADEMARK OFFICE

SIGNATURE

NAME STEPHEN B. MAEBIUS

REGISTRATION NUMBER 35,264

#6

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Atty. Docket No: 065691/0270

In re patent application of

PITIOT, GILLES *et al.*

Serial No. 10/049,372

Filed: February 12, 2002

For: HUMAN ODORANT-BINDING PROTEINS WHICH BIND HYDROPHOBIC LIGANDS: POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAID POLYPEPTIDES, AND USES THEREOF

STATEMENT TO SUPPORT FILING AND SUBMISSION IN ACCORDANCE WITH 37 C.F.R. §§ 1.821-1.825

Assistant Commissioner for Patents
Washington, D.C. 20231
Box SEQUENCE

Sir:

In connection with a Sequence Listing submitted concurrently herewith, the undersigned hereby states that:

1. the submission, filed herewith in accordance with 37 C.F.R. § 1.821(g), does not include new matter;

2. the content of the attached paper copy and the attached computer readable copy of the Sequence Listing, submitted in accordance with 37 C.F.R. § 1.821(c) and (e), respectively, are the same; and

3. all statements made herein of their own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United

States Code and that such willful false statements may jeopardize the validity of the application or any patent resulting therefrom.

Respectfully submitted,

August 6, 2002
Date

James A. Coburn

HARBOR CONSULTING
Intellectual Property Services
1500A Lafayette Road
Suite 262
Portsmouth, N.H.
800-318-3021

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No. :

U.S. National Serial No. :

Filed :

PCT International Application No. : PCT/FR00/02319

VERIFICATION OF A TRANSLATION

I, the below named translator, hereby declare that:

My name and post office address are as stated below;

That I am knowledgeable in the French language in which the below identified international application was filed, and that, to the best of my knowledge and belief, the English translation of the amended sheets of the international application No. PCT/FR00/02319 is a true and complete translation of the amended sheets of the above identified international application as filed.

I hereby declare that all the statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the patent application issued thereon.

Date: February 28, 2002

Full name of the translator : Elaine Patricia PARRISH

For and on behalf of RWS Group plc

Post Office Address : Europa House, Marsham Way,

Gerrards Cross, Buckinghamshire,

England.

#6

Atty. Dkt. No. 065691/0270

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: Gilles PITIOT et al.

Title: HUMAN ODORANT-BINDING PROTEINS WHICH BIND HYDROPHOBIC LIGANDS: POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAID POLYPEPTIDES, AND USES THEREOF

Appl. No. 10/049,372

Filing Date: 08/15/2002

Examiner: Unassigned

Art Unit: Unassigned

AMENDMENT IN RESPONSE TO NOTICE UNDER 37 CFR §§1.821-825

Commissioner for Patents
Washington, D.C. 20231

Sir:

In response to the Notice to Comply With Requirements for Applications Containing Sequence Disclosures mailed April 16, 2002, please amend the application as follows:

In the Specification:

Please amend the specification as shown:

Please delete the paragraph on page 6, lines 28-30, and replace it with the following paragraph:

The polypeptides according to the invention are characterized in that they preferably comprise the Gly-Thr-Trp-Tyr (SEQ ID NO: 25) domain.

Please delete the paragraph on page 33, lines 6-15, and replace it with the following paragraph:

The upper lines represent the hOBPIIa sequences (SEQ ID NOS 17-19, 21-24) and the lower lines represent the hOBPIIb sequence (SEQ ID NO: 20) for which only the different nucleotides are represented; a dash indicates the absence of corresponding sequences. The shaded capitals are the exon sequences and the lower case letters are the intron sequences. The sizes indicated on the left are indicated in bp. The TATA box is in bold characters and the polyadenylation signal is underlined. The boxes indicate the splice acceptor sites for exons 5, 5b and 5c.

Please delete the paragraph on page 33, lines 35-39 thru page 34 lines 1-4, and replace it with the following paragraph:

FIGURE 5: Aligment of the protein sequences derived from the two human genes hOBPIIa and hOBPIIb (hOBPIIaα = OBP2aaHOMSA (SEQ ID NO: 2), hOBPIIbα = OBP2baHOMSA (SEQ ID NO: 10), hOBPIIbβ = OPB2bbHOMSA (residues 15-179 of SEQ ID NO: 12), hOBPIIaγ = OBP2agHOMSA (SEQ ID NO: 6), hOBPIIaβ = OBP2abHOMSA (SEQ ID NO: 4)), of human tear lipocalins (LCN1_HOMSA)(SEQ ID NO: 46), of rat OBPII (OBP2_RATNO)(SEQ ID NO: 47), of bovine lactoglobulin BLG (LACB_BOSTA)(SEQ ID NO: 48), of mouse MUP (MUP6_MUSMU)(SEQ ID NO: 49), of human RBP (RBP_HOMSA)(SEQ ID NO: 50), of bovine OBP (OBP_BOSTA)(SEQ ID NO: 51), of rat MUP (MUP_RATNO)(SEQ ID NO: 52) and of porcine OBP (OBP_SUSSC)(SEQ ID NO: 53).

Please delete the paragraph on page 34, lines 6-16 and replace it with the following paragraph:

The residues in the dark grey boxes are identical and those in the light gray boxes are similar. The elements of secondary structure predicted with the DSC program are underlined and the amino acid residues are in italics. The β-sheets and the α-helices are

numbered for hOPBIIa and b. The anticipated signal peptide cleavage site is indicated with an arrow (AAA↓LS) at position 15. Nonaligned sequences of very divergent forms of spliced genes, hOPBaδ (OPB2adHOMSA)(SEQ ID NO: 8) and hOBPbγ (OBP2bgHOMSA)(SEQ ID NO: 14) have been added at the bottom of the alignment after the analysis.

Please delete the paragraph on page 39, lines 30-38 thru page 40, lines 1-2, and replace it with the following paragraph:

A Lambda gt11 human testes cDNA library (Clontech) (10^7 p.f.u.) was amplified by 30 polymerase chain reaction (PCR) cycles (94°C 45 sec, 54°C 45 sec, 72°C 1 min 30 sec) with the primer oliEST58 CCTGCAGGTACATGAGCTTCC (SEQ ID NO: 26) and 5′ or 3′ screening amplimers for inserts located on the arms of the lambda gt11 vectors. A nested PCR was then carried out with oliEST26 CGCTGTATTTGCCAGGCTCC (SEQ ID NO: 27) and oligonucleotides specific for the arm of the vector. The PCR products were subcloned into the pGEM-T (r) vector, which made it possible to obtain the 5′ end of the hOBPII gene cDNAs.

Please delete the paragraph on page 40, lines 20-39 thru page 41, lines 1-5, and replace it with the following paragraph:

Tissue samples were collected from 45 to 55 year old Caucasian individuals, in agreement with the French regulations in force. The total RNA is extracted according to a single-step method using the RNA NOW® reagent, according to the manufacturer's (Biogentex) instructions. 5 µg of total RNA were reversed-transcribed in a final volume of 20 µl using 0.5 ng of oligonucleotide GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT (SEQ ID NO: 28) with the Superscript® preamplification system (Gibco BRL). Three µl of these reactions were then used for the following PCRs. The expression of the specific mRNAs was determined by PCR using: the primers TL: CCTCTCCCAGCCCCAGCAAG (SEQ ID NO: 29) and AP: GACTCGAGTCGACATCG (SEQ

ID NO: 30) for the LCN1-type genes (LCN1, LCN1b, LCN1c) and, for the hOBPII-type genes, the primers DE: CGCCCAGTGACCTGCCGAGGTC (SEQ ID NO: 31), and FI: CTTTATTTGGAGTCAGGTGGGTG (SEQ ID NO: 32). As a quality control for the RNA's the primers G3PDH1: CTCTGCCCCCTCTGCTGATG (SEQ ID NO: 33) and G3PDH2: CCTGCTTCACCACCTTCTTG (SEQ ID NO: 34) for the G3PDH gene were used; the G3PDH gene is considered to be constitutively expressed in all cell types. 32 PCR cycles (94°C 45 sec, 54°C 45 sec, 72°C 2 min 30 sec) were performed and the amplification products were separated on 1% agarose gel. The DNA was transferred onto a Hybond N+® membrane.

Please delete the paragraph on page 41, lines 7-16, and replace it with the following paragraph:

For the detection of the expression of the various genes, several oligonucleotides specific for the respective genes were synthesized:

- olLCN1: GACTCAGACTCCGGAGATGA (SEQ ID NO: 35),

- olLCN1b: AACTCAGACACCAGAGATGA (SEQ ID NO: 36),

- olLCN1c: GACTCAGATCCCGGAGATGA (SEQ ID NO: 37),

- ol5: CCAGGAGGGACCACTACA (SEQ ID NO: 38) specific for the hOBPIIb gene,

- ol4: CCGGGACGGACGACTACG (SEQ ID NO: 39) specific for the hOBPIIa gene,

- G3PDH3: CTCATGACCACAGTCCATGC (SEQ ID NO: 40).

Please delete the paragraph on page 41, lines 30-39 thru page 42, lines1-7, and replace it with the following paragraph:

The genotyping is carried out with PCR reactions using 100 ng of genomic DNA originating from 8 CEPH reference families, and using the oligonucleotides oli9 TGTTCGGGAACGCAGCTT (SEQ ID NO: 41) and

oli10a TGCCGCTGTCCCCACGTCGG (SEQ ID NO: 42). The thermocycler parameters consist of an initial cycle at 94°C for 10 min followed by 30 cycles at 94°C for 30 sec, 55°C for 30 sec and 70°C for 45 sec, and then a final elongation step of 10 min at 70°C. The PCR products are then analyzed on a 3% agarose gel. The information regarding the markers for chromosome 9 may be obtained at the following internet address (hTTP://galton.ucl.ac.uk); the analyses were carried out using the linkage study tools previously described in Lacazette *et al.* (1997). The haplotypes are reconstructed manually according to the recombination events previously described in family 1362 (Attwood *et al.*, 1994).

Please delete the paragraph on page 52, lines 34-39 thru page 54, lines1-11, and replace it with the following paragraph:

In the case of the hOBPIIb gene, in addition to the hOBPIIbα mRNA previously described, a 106 bp super-numerary exon (exon 3b) between the previous exons 3 and 4 was identified (figure 3). This longer mRNA (782 nucleotides) encodes a 165 amino acid protein hOBPIIbβ. From a protein structure point of view, hOBPIIbβ is identical to hOBPIIbα up to the 5^{th} putative β-sheet and then differs due to a reading frame shift. The predictions from the computer programs indicate that the motif ALWEALAIDTRLK (SEQ ID NO: 43) is an α-helix which is just behind the fifth β-sheet. Two additional β-sheets may be present in the long C-terminal portion.

Please delete the paragraph on page 58, lines 36-39 thru page 59, lines1-23, and replace it with the following paragraph:

A PCR on the plasmid DNA of an hOBPIIbα clone using the primers BIIa/b (5' GTC GGA TCC CTG TCC TTC ACC CTG GAG G 3') (SEQ ID NO: 44), a sense oligonucleotide beginning 45 bases after the protein-initiating ATG, and XIIb (5; GTC CTC GAG GTG TTC GGG AAC GCA GCT TC 3') (SEQ ID NO: 45), an antisense oligonucleotide preceding the stop

codon of the hOBPIIbα protein, made it possible to amplify all of the DNA encoding the secreted hOBPIIbα protein. The BanH I and Xho I enzymatic restriction sites located at the ends of the two oligonucleotides (bases underlined) were used for directional cloning into a plasmid expression vector pGEX-6Pl, followed by transformation by electroporation (1800 V, 200 Ω, 25 µF) into bacterial strain BL21. Synthesis of the recombinant protein is obtained by adding IPTG, at 5 mM final for 3 h, to 250 ml of culture of the strain in LB medium containing 100 µg/ml of ampicillin pre-incubated at 37°C for 2 h. The centrifuged cultures are taken up in 25 ml of TENGN buffer. The lysate is sonicated and recentrifuged. The fusion protein is then purified using 4 ml of beads bearing covalently attached insoluble glutathione (Sigma) per 25 ml of supernatant. After incubation for 4 h, they are washed with 3 volumes of 1M NaCl and then with 10 volumes of 1 X PBS. The elution is obtained by bringing the beads into contact with a glutathione solution (10 mM reduced glutathione, 50 mM Tris-HCl, pH 8.0) for 10 min.

REMARKS

Applicants believe that the present application is now in condition for allowance. Favorable reconsideration of the application as amended is respectfully requested.

The Examiner is invited to contact the undersigned by telephone if it is felt that a telephone interview would advance the prosecution of the present application.

Respectfully submitted,

Date August 14, 2002

By _____

FOLEY & LARDNER
Customer Number: 22428
22428
PATENT TRADEMARK OFFICE
Telephone: (202) 672-5407
Facsimile: (202) 672-5399

Stephen Maebius
Attorney for Applicants
Registration No. 35,264

Should additional fees be necessary in connection with the filing of this paper, or if a petition for extension of time is required for timely acceptance of same, the Commissioner is hereby authorized to charge Deposit Account No.19-0741 for any such fees; and applicant(s) hereby petition for any needed extension of time.

MARKED UP VERSION ATTACHED TO AMENDMENT IN

SERIAL NO. 10/049,372

Marked up version of the paragraph on page 6, lines 28-30 is below:

The polypeptides according to the invention are characterized in that they preferably comprise the Gly-Thr-Trp-Tyr **(SEQ ID NO: 25)** domain.

Marked up version of the paragraph on page 33, lines 6-15, is below:

The upper lines represent the hOBPIIa sequences **(SEQ ID NOS 17-19, 21-24)** and the lower lines represent the hOBPIIb sequence **(SEQ ID NO: 20)** for which only the different nucleotides are represented; a dash indicates the absence of corresponding sequences. The shaded capitals are the exon sequences and the lower case letters are the intron sequences. The sizes indicated on the left are indicated in bp. The TATA box is in bold characters and the polyadenylation signal is underlined. The boxes indicate the splice acceptor sites for exons 5, 5b and 5c.

Marked up version of the paragraph on page 33, lines 35-39 thru page 34 lines 1-4, is below:

FIGURE 5: Aligment of the protein sequences derived from the two human genes hOBPIIa and hOBPIIb (hOBPIIaα = OBP2aaHOMSA (SEQ ID NO: 2), hOBPIIbα = OBP2baHOMSA (SEQ ID NO: 10), hOBPIIbβ = OPB2bbHOMSA (residues 15-179 of SEQ ID NO: 12), hOBPIIaγ = OBP2agHOMSA (SEQ ID NO: 6), hOBPIIaβ = OBP2abHOMSA (SEQ ID NO: 4)), of human tear lipocalins (LCN1_HOMSA) (SEQ ID NO: 46), of rat OBPII (OBP2_RATNO) (SEQ ID NO: 47), of bovine lactoglobulin BLG (LACB_BOSTA) (SEQ ID NO: 48), of mouse MUP (MUP6_MUSMU) (SEQ ID NO: 49), of human RBP (RBP_HOMSA) (SEQ ID NO: 50), of bovine OBP (OBP_BOSTA) (SEQ ID NO: 51), of rat MUP (MUP_RATNO) (SEQ ID NO: 52) and of porcine OBP (OBP_SUSSC) (SEQ ID NO: 53).

Marked up version of the paragraph on page 34, lines 6-16, is below:

The residues in the dark grey boxes are identical and those in the light gray boxes are similar. The elements of secondary structure predicted with the DSC program are underlined and the amino acid residues are in italics. The β-sheets and the α-helices are numbered for hOPBIIa and b. The anticipated signal peptide cleavage site is indicated with an arrow (AAA↓LS) at position 15. Nonaligned sequences of very divergent forms of spliced genes, hOPBaδ (OPB2adHOMSA) **(SEQ ID NO: 8)** and hOBPbγ (OBP2bgHOMSA) **(SEQ ID NO: 14)** have been added at the bottom of the alignment after the analysis.

Marked up version of the paragraph on page 39, lines 30-38 thru page 40, lines 1-2, is below:

A Lambda gt11 human testes cDNA library (Clontech) (10^7 p.f.u.) was amplified by 30 polymerase chain reaction (PCR) cycles (94°C 45 sec, 54°C 45 sec, 72°C 1 min 30 sec) with the primer oliEST58 CCTGCAGGTACATGAGCTTCC **(SEQ ID NO: 26)** and 5′ or 3′ screening amplimers for inserts located on the arms of the lambda gt11 vectors. A nested PCR was then carried out with oliEST26 CGCTGTATTTGCCAGGCTCC **(SEQ ID NO: 27)** and oligonucleotides specific for the arm of the vector. The PCR products were subcloned into the pGEM-T (r) vector, which made it possible to obtain the 5′ end of the hOBPII gene cDNAs.

Marked up version of the paragraph on page 40, lines 20-39 thru page 41, lines 1-5, is below:

Tissue samples were collected from 45 to 55 year old Caucasian individuals, in agreement with the French regulations in force. The total RNA is extracted according to a single-step method using the

RNA NOW® reagent, according to the manufacturer's (Biogentex) instructions. 5 µg of total RNA were reversed-transcribed in a final volume of 20 µl using 0.5 ng of oligonucleotide GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT **(SEQ ID NO: 28)** with the Superscript® preamplification system (Gibco BRL). Three µl of these reactions were then used for the following PCRs. The expression of the specific mRNAs was determined by PCR using: the primers TL: CCTCTCCCAGCCCCAGCAAG **(SEQ ID NO: 29)** and AP: GACTCGAGTCGACATCG **(SEQ ID NO: 30)** for the LCN1-type genes (LCN1, LCN1b, LCN1c) and, for the hOBPII-type genes, the primers DE: CGCCCAGTGACCTGCCGAGGTC **(SEQ ID NO: 31)**, and FI: CTTTATTTGGAGTCAGGTGGGTG **(SEQ ID NO: 32)**. As a quality control for the RNA's the primers G3PDH1: CTCTGCCCCCTCTGCTGATG **(SEQ ID NO: 33)** and G3PDH2: CCTGCTTCACCACCTTCTTG **(SEQ ID NO: 34)** for the G3PDH gene were used; the G3PDH gene is considered to be constitutively expressed in all cell types. 32 PCR cycles (94°C 45 sec, 54°C 45 sec, 72°C 2 min 30 sec) were performed and the amplification products were separated on 1% agarose gel. The DNA was transferred onto a Hybond N+® membrane.

Marked up version of the paragraph on page 41, lines 6-16, is below:

For the detection of the expression of the various genes, several oligonucleotides specific for the respective genes were synthesized:

- olLCN1 : GACTCAGACTCCGGAGATGA **(SEQ ID NO: 35)**,

- olLCN1b: AACTCAGACACCAGAGATGA **(SEQ ID NO: 36)**,

- olLCN1c: GACTCAGATCCCGGAGATGA **(SEQ ID NO: 37)**,

- ol5: CCAGGAGGGACCACTACA **(SEQ ID NO: 38)** specific for the hOBPIIb gene,

- ol4: CCGGGACGGACGACTACG **(SEQ ID NO: 39)** specific for the hOBPIIa gene,

- G3PDH3: CTCATGACCACAGTCCATGC **(SEQ ID NO: 40)**.

Marked up version of the paragraph on page 41, lines 30-39 thru page 42, lines1-7, is below:

The genotyping is carried out with PCR reactions using 100 ng of genomic DNA originating from 8 CEPH reference families, and using the oligonucleotides oli9 TGTTCGGGAACGCAGCTT **(SEQ ID NO: 41)** and oli10a TGCCGCTGTCCCCACGTCGG **(SEQ ID NO: 42)**. The thermocycler parameters consist of an initial cycle at 94°C for 10 min followed by 30 cycles at 94°C for 30 sec, 55°C for 30 sec and 70°C for 45 sec, and then a final elongation step of 10 min at 70°C. The PCR products are then analyzed on a 3% agarose gel. The information regarding the markers for chromosome 9 may be obtained at the following internet address (hTTP://galton.ucl.ac.uk); the analyses were carried out using the linkage study tools previously described in Lacazette *et al.* (1997). The haplotypes are reconstructed manually according to the recombination events previously described in family 1362 (Attwood *et al.*, 1994).

Marked up version of the paragraph on page 52, lines 34-39 thru page 54, lines1-11, is below:

In the case of the hOBPIIb gene, in addition to the hOBPIIbα mRNA previously described, a 106 bp super-numerary exon (exon 3b) between the previous exons 3 and 4 was identified (figure 3). This longer mRNA (782 nucleotides) encodes a 165 amino acid protein hOBPIIbβ. From a protein structure point of view, hOBPIIbβ is identical to hOBPIIbα up to the 5^{th} putative β-sheet and then differs due to a reading frame shift. The predictions from the computer programs indicate that the motif ALWEALAIDTRLK **(SEQ ID NO: 43)** is an α-helix which is just behind the fifth β-sheet. Two additional β-sheets may be present in the long C-terminal portion.

Marked up version of the paragraph on page 58, lines 36-39 thru page 59, lines1-23, is below:

A PCR on the plasmid DNA of an hOBPIIbα clone using the primers BIIa/b (5′ GTC GGA TCC CTG TCC TTC ACC CTG GAG G 3′) **(SEQ ID NO: 44)**, a sense oligonucleotide beginning 45 bases after the protein-initiating ATG, and XIIb (5; GTC CTC GAG GTG TTC GGG AAC GCA GCT TC 3′) **(SEQ ID NO: 45)**, an antisense oligonucleotide preceding the stop codon of the hOBPIIbα protein, made it possible to amplify all of the DNA encoding the secreted hOBPIIbα protein. The BanH I and Xho I enzymatic restriction sites located at the ends of the two oligonucleotides (bases underlined) were used for directional cloning into a plasmid expression vector pGEX-6Pl, followed by transformation by electroporation (1800 V, 200 Ω, 25 µF) into bacterial strain BL21. Synthesis of the recombinant protein is obtained by adding IPTG, at 5 mM final for 3 h, to 250 ml of culture of the strain in LB medium containing 100 µg/ml of ampicillin pre-incubated at 37°C for 2 h. The centrifuged cultures are taken up in 25 ml of TENGN buffer. The lysate is sonicated and recentrifuged. The fusion protein is then purified using 4 ml of beads bearing covalently attached insoluble glutathione (Sigma) per 25 ml of supernatant. After incubation for 4 h, they are washed with 3 volumes of 1M NaCl and then with 10 volumes of 1 X PBS. The elution is obtained by bringing the beads into contact with a glutathione solution (10 mM reduced glutathione, 50 mM Tris-HCl, pH 8.0) for 10 min.

CLAIMS

1. An isolated polypeptide comprising an amino acid sequence having at least 90% identity with the amino acid sequences SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 12 or SEQ ID No. 14.

2. An isolated polypeptide, characterized in that it comprises a polypeptide chosen from:

a) a polypeptide of sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 12 or SEQ ID No. 14;

b) a polypeptide which is a variant of a polypeptide of amino acid sequences defined in a);

c) a polypeptide homologous to the polypeptide defined in a) or b) and comprising at least 90% identity with said polypeptide of a);

d) a fragment of at least 15 consecutive amino acids of a polypeptide defined in a), b) or c);

e) a biologically active fragment of a polypeptide defined in a), b) or c).

3. An isolated polypeptide selected from a polypeptide corresponding to the sequence SEQ ID No. 2 and named $OBPII_{a\alpha}$, to the sequence SEQ ID No. 4 and named $OBPII_{a\beta}$, to the sequence SEQ ID No. 6 and named $OBPII_{a\gamma}$, and to the sequence SEQ ID No. 12 and named $OBPII_{b\beta}$.

4. The polypeptide as claimed in any one of claims 1 to 3, characterized in that it comprises at least the Gly-Thr-Trp-Tyr domain.

5. An isolated polynucleotide, characterized in that it encodes a polypeptide as claimed in any one of claims 1 to 4.

6. The polynucleotide as claimed in claim 5, of sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 11 or SEQ ID No. 13.

7. An isolated polynucleotide, characterized in that it comprises a polynucleotide chosen from:

a) a polynucleotide of sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 11 or SEQ ID No. 13, or the sequence of which is that of the RNA corresponding to the sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 11 or SEQ ID No. 13;

b) a polynucleotide, the sequence of which is complementary to the sequence of a polynucleotide defined in a),

c) a polynucleotide, the sequence of which comprises at least 80% identity with a polynucleotide defined in a) or b),

d) a fragment of at least 15 consecutive nucleotides, in preference 21 consecutive nucleotides, and preferably 30 consecutive nucleotides, of a polynucleotide defined in a), b), c) or d).

8. The polynucleotide as claimed in claim 7, characterized in that it is labeled directly or indirectly with a radioactive compound or a nonradioactive compound.

9. The use of a polynucleotide as claimed in claim 8, as a probe for detecting nucleic acid sequences.

10. The use of a polynucleotide as claimed in any one of claims 5 to 8, as a primer for amplifying or polymerizing nucleic acid sequences.

11. The use of a polynucleotide as claimed in claim 8, as a sense or antisense oligonucleotide for controlling the expression of the corresponding protein product.

12. A recombinant vector for cloning a polynucleotide as claimed in one of claims 5 to 8 and/or for expressing a polypeptide as claimed in one of claims 1 to 4, characterized in that it contains a polynucleotide as claimed in any one of claims 5 to 8.

13. The vector as claimed in claim 12, characterized in that it comprises the elements which allow the expression, optionally the secretion, of said polypeptide in a host cell.

14. The vector as claimed in any one of claims 12 to 13, characterized in that the elements which allow the expression of said polypeptide are chosen from:

a) the isolated polynucleotide of sequence SEQ ID No. 15 and SEQ ID No. 16;

b) a polynucleotide, the sequence of which is complementary to the sequence of the polynucleotide defined in a);

c) a polynucleotide, the sequence of which comprises at least 80% identity with a polynucleotide defined in a) or in b);

d) a polynucleotide which hybridizes, under conditions of high stringency, with a sequence of the polynucleotide defined in a), b) or c).

15. The vector as claimed in claims 13 and 14, for expression in eukaryotic cells, selected from viral DNA and naked DNA.

16. A host cell, characterized in that it is transformed with a vector as claimed in one of claims 12 to 15.

17. A process for preparing a recombinant polypeptide, characterized in that a host cell as claimed in claim 16 is cultured under conditions which allow the expression and, optionally, the secretion of said recombinant polypeptide, and in that said recombinant polypeptide is recovered.

18. A recombinant polypeptide which can be obtained using a process as claimed in claim 17.

19. The use of a polypeptide chosen from a polypeptide as claimed in any one of claims 1 to 4 and 18, or a fragment thereof, as a protein which binds to a hydrophobic ligand, preferably an odorous molecule.

20. The use of a polypeptide as claimed in any one of claims 1 to 4 and 18, or of the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$, a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$ or a fragment thereof, as a competitive inhibitor, as an agonist or as an antagonist of the cellular receptors for lipocalins.

21. A monoclonal or polyclonal antibody, and fragments thereof, characterized in that it is specifically against an isolated polypeptide as claimed in one of claims 1 to 4 and 18.

22. The use of an antibody as claimed in claim 21, for demonstrating the presence of a polypeptide as claimed in one of claims 1 to 4 and 18 in a biological sample.

23. A process for detecting an antibody against hOBPII in human serum from an allergic and/or asthmatic patient, using an hOBPII polypeptide.

24. A process for controlling the volatilization of an odorant, characterized in that it comprises a step of binding of said odorant with a polypeptide as claimed in any one of claims 1 to 4 and 18, or with the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$.

25. The process as claimed in claim 24, characterized in that the polypeptide is bound to a solid support.

26. The process as claimed in claim 24, characterized in that the polypeptide is in a liquid composition.

27. The process as claimed in claim 26, characterized in that said composition is a fragranced composition for the skin.

28. A process for screening a molecule, preferably odorants or flavors, which comprises passing the molecule over a substrate which comprises a polypeptide as claimed in any one of claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, bound to said substrate, said polypeptide

binding said odorant or flavor, and recovering said odorant or flavor from the polypeptide if necessary.

29. The process as claimed in claim 28, characterized in that the molecules are human pheromones.

30. A process for solubilizing lipophilic molecules, characterized in that it comprises binding said molecule to a polypeptide as claimed in any one of claims 1 to 4 and 18, or to the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$.

31. The application of the polypeptides as claimed in one of claims 1 to 4 and 18, or of the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, of a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or of a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, in combination with dietary fatty acids, as a food supplement.

32. The application as claimed in claim 31, for preparing a medicinal product intended for the treatment of hyperlipidemias and obesity.

33. The application as claimed in claim 31, for supplementing nonmaternal milks.

34. The polypeptide as claimed in any one of claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, as an agent for targeting a pharmaceutical compound.

35. The polypeptide as claimed in claim 34, characterized in that said polypeptide is expressed in the form of a protein from fusion with a protein which allows specific cellular addressing.

36. The polypeptide as claimed in claim 35, characterized in that said protein which allows specific cellular addressing is chosen from the group composed of interleukins, of cytokines, of lymphokines, of interferons, of growth factors, of hormones and of antibodies.

37. The polypeptide as claimed in claim 34, characterized in that said polypeptide is combined with a molecule which allows specific cellular addressing.

38. The polypeptide as claimed in claim 37, characterized in that said molecule which allows specific cellular addressing is chosen from the group composed of steroids, of interleukins, of cytokines, of lymphokines, of interferons, of growth factors, of hormones and of antibodies.

39. The polypeptide as claimed in any one of claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, as a pharmaceutical compound transporter.

40. A pharmaceutical composition comprising a pharmaceutical compound bound at least to a polypeptide as claimed in one of claims 34 to 39, and a pharmaceutically acceptable vehicle.

41. The pharmaceutical composition as claimed in

claim 40, characterized in that the pharmaceutical compound is chosen from the group of anticancer agents.

42. The pharmaceutical composition as claimed in claim 41, characterized in that said anticancer agent is a radioactive isotope chosen from the group: iodine^{131}, yttrium^{90}, gold^{199}, palladium^{100}, copper^{67}, bismuth^{217} and antimony^{211}.

43. The pharmaceutical composition as claimed in any one of claims 40 to 42, characterized in that said polypeptide as claimed in any one of claims 34 to 39 constitutes a delayed form of delivery of said pharmaceutical compound in the body.

44. A pharmaceutical composition comprising an expression vector as claimed in claim 12 or 13, and a pharmaceutically acceptable vehicle.

45. The pharmaceutical composition as claimed in any one of claims 40 to 44, for treating cancer preferably chosen from breast cancer, uterine cancer, prostate cancer, liver cancer and pulmonary epithelial cell carcinoma.

46. The use of a polypeptide chosen from the polypeptide hOBPII$_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide hOBPII$_{b\alpha}$ and a polypeptide homologous to said polypeptide hOBPII$_{b\alpha}$ comprising at least 90% identity with said polypeptide hOBPII$_{b\alpha}$, for preparing a medicinal product intended for the treatment of uterine cancer, prostate cancer, liver cancer and pulmonary epithelial cell cancer.

47. The use of the polypeptide as claimed in any one of claims 1 to 4 and 18, or the polypeptide hOBPII$_{b\alpha}$ of sequence SEQ ID No. 10, or a polypeptide which is a

variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, for preparing a medicinal product intended for the transport of a compound across the placental barrier.

48. The use as claimed in claim 47, characterized in that said transport is carried out from the gestating mother to the fetus, and in that said compound is chosen from hormones, essential fatty acids, lipophilic medicinal products and vitamins.

49. The use as claimed in claim 47, characterized in that said transport is carried out from the fetus to the mother and is intended for the detoxification of the fetus.

50. The polypeptide as claimed in claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, as a pregnancy marker.

51. The polypeptide as claimed in claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, as a marker for a feto-placental pathological condition

SEQUENCE LISTING

<110> PITIOT, GILLES
LACAZETTE, ERIC
GACHON, FRANCOISE

<120> HUMAN ODORANT-BINDING PROTEINS WHICH BIND HYDROPHOBIC LIGANDS: POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAID POLYPEPTIDES, AND USES THEREOF

<130> 065691/0270

<140> 10/049,372
<141> 2002-02-12

<150> PCT/FR00/02319
<151> 2000-08-11

<150> FR 99 10439
<151> 1999-08-12

<160> 53

<170> PatentIn Ver. 2.1

<210> 1
<211> 676
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(552)

<220>
<223> cDNA396 (676) /gl (hOBPIIa-alpha)

<400> 1
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
Met Lys Thr Leu
1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg 102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
5 10 15 20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc 150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
25 30 35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg 198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
40 45 50

aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc 246
Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
55 60 65

```
atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg   294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
     70                  75                  80

gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac   342
Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr
 85                  90                  95                 100

ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac   390
Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
                105                 110                 115

cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg ggt agg aat   438
Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Gly Arg Asn
            120                 125                 130

cct aat acc aac ctg gag gcc ctg gaa gaa ttt aag aaa ttg gtg cag   486
Pro Asn Thr Asn Leu Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln
        135                 140                 145

cac aag gga ctc tcg gag gag gac att ttc atg ccc ctg cag acg gga   534
His Lys Gly Leu Ser Glu Glu Asp Ile Phe Met Pro Leu Gln Thr Gly
    150                 155                 160

agc tgc gtt ctc gaa cac taggcagccc ccgggtctgc acctccagag          582
Ser Cys Val Leu Glu His
165                 170

cccaccctac caccagacac agagcccgga ccacctggac ctaccctcca gccatgaccc 642

ttccctgctc ccacccacct gactccaaat aaag                             676
```

<210> 2
<211> 170
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
  1               5                  10                  15

Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
             20                  25                  30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
         35                  40                  45

Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
     50                  55                  60

Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65                  70                  75                  80

Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
                 85                  90                  95
```

```
Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
            100                 105                 110

Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
        115                 120                 125

Val Gly Arg Asn Pro Asn Thr Asn Leu Glu Ala Leu Glu Glu Phe Lys
    130                 135                 140

Lys Leu Val Gln His Lys Gly Leu Ser Glu Glu Asp Ile Phe Met Pro
145                 150                 155                 160

Leu Gln Thr Gly Ser Cys Val Leu Glu His
                165                 170
```

<210> 3
<211> 725
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(480)

<220>
<223> cDNA396 (725) /SM12 (hOBPIIa-beta)

<400> 3

```
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg    54
                                                  Met Lys Thr Leu
                                                    1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg   102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
  5                  10                  15                  20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc   150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
                25                  30                  35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg   198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
            40                  45                  50

aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc   246
Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
        55                  60                  65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg   294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
    70                  75                  80

gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac   342
Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr
 85                  90                  95                 100
```

```
ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac   390
Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
                105                 110                 115

cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg ggg ccg tgc   438
Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Gly Pro Cys
            120                 125                 130

cgc tgt ccc cac gtc ggc tca cct ggc cac ctc acc tgc agg           480
Arg Cys Pro His Val Gly Ser Pro Gly His Leu Thr Cys Arg
        135                 140                 145

taggaatcct aataccaacc tggaggccct ggaagaattt aagaaattgg tgcagcacaa 540

gggactctcg gaggaggaca ttttcatgcc cctgcagacg ggaagctgcg ttctcgaaca 600

ctaggcagcc cccgggtctg cacctccaga gcccacccta ccaccagaca cagagcccgg 660

accacctgga cctaccctcc agccatgacc cttccctgct cccacccacc tgactccaaa 720

taaag                                                             725
```

<210> 4
<211> 146
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
  1               5                  10                  15

Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
            20                  25                  30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
        35                  40                  45

Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
    50                  55                  60

Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
65                  70                  75                  80

Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
                85                  90                  95

Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
            100                 105                 110

Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
        115                 120                 125

Val Gly Pro Cys Arg Cys Pro His Val Gly Ser Pro Gly His Leu Thr
    130                 135                 140

Cys Arg
145
```

<210> 5
<211> 741
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(726)

<220>
<223> cDNA396 (741) /SM4 (hOBPIIa-gamma)

<400> 5

```
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg    54
                                                   Met Lys Thr Leu
                                                     1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg   102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
  5                  10                  15                  20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc   150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
                 25                  30                  35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg   198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
             40                  45                  50

aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc   246
Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
         55                  60                  65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg   294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
     70                  75                  80

gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac   342
Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr
 85                  90                  95                 100

ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac   390
Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
                105                 110                 115

cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg gca tct gct   438
Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Ala Ser Ala
            120                 125                 130

ccc tgc agg gcc gtg ccg ctg tcc cca cgt cgg ctc acc tgg cca cct   486
Pro Cys Arg Ala Val Pro Leu Ser Pro Arg Arg Leu Thr Trp Pro Pro
        135                 140                 145
```

```
cac ctg cag gta gga atc cta ata cca acc tgg agg ccc tgg aag aat   534
His Leu Gln Val Gly Ile Leu Ile Pro Thr Trp Arg Pro Trp Lys Asn
    150                 155                 160

tta aga aat tgg tgc agc aca agg gac tct cgg agg agg aca ttt tca   582
Leu Arg Asn Trp Cys Ser Thr Arg Asp Ser Arg Arg Arg Thr Phe Ser
165                 170                 175                 180

tgc ccc tgc aga cgg gaa gct gcg ttc tcg aac act agg cag ccc ccg   630
Cys Pro Cys Arg Arg Glu Ala Ala Phe Ser Asn Thr Arg Gln Pro Pro
                185                 190                 195

ggt ctg cac ctc cag agc cca ccc tac cac cag aca cag agc ccg gac   678
Gly Leu His Leu Gln Ser Pro Pro Tyr His Gln Thr Gln Ser Pro Asp
            200                 205                 210

cac ctg gac cta ccc tcc agc cat gac cct tcc ctg ctc cca ccc acc   726
His Leu Asp Leu Pro Ser Ser His Asp Pro Ser Leu Leu Pro Pro Thr
        215                 220                 225

tgactccaaa taaag                                                  741
```

<210> 6
<211> 228
<212> PRT
<213> Homo sapiens

<400> 6

```
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
  1               5                  10                  15

Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
             20                  25                  30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
         35                  40                  45

Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
     50                  55                  60

Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65                  70                  75                  80

Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
                 85                  90                  95

Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
            100                 105                 110

Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
        115                 120                 125

Val Ala Ser Ala Pro Cys Arg Ala Val Pro Leu Ser Pro Arg Arg Leu
    130                 135                 140

Thr Trp Pro Pro His Leu Gln Val Gly Ile Leu Ile Pro Thr Trp Arg
145                 150                 155                 160
```

```
Pro Trp Lys Asn Leu Arg Asn Trp Cys Ser Thr Arg Asp Ser Arg Arg
                165                 170                 175

Arg Thr Phe Ser Cys Pro Cys Arg Arg Glu Ala Ala Phe Ser Asn Thr
            180                 185                 190

Arg Gln Pro Pro Gly Leu His Leu Gln Ser Pro Pro Tyr His Gln Thr
        195                 200                 205

Gln Ser Pro Asp His Leu Asp Leu Pro Ser Ser His Asp Pro Ser Leu
    210                 215                 220

Leu Pro Pro Thr
225
```

<210> 7
<211> 607
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(483)

<220>
<223> cDNA396 (607) - short form (hOBPIIa-delta)

<400> 7

```
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg    54
                                                  Met Lys Thr Leu
                                                    1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg   102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
  5                  10                  15                  20

gag gag gag gat gag gga gaa tcg gtg cat cca gaa gaa aat cct gat   150
Glu Glu Glu Asp Glu Gly Glu Ser Val His Pro Glu Glu Asn Pro Asp
                25                  30                  35

gcg gaa gac gga gga gcc tgg caa att cag cgc cta tgg ggg cag gaa   198
Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu Trp Gly Gln Glu
            40                  45                  50

gct cat ata cct gca gga gct gcc cgg gac gga cga cta cgt ctt tta   246
Ala His Ile Pro Ala Gly Ala Ala Arg Asp Gly Arg Leu Arg Leu Leu
        55                  60                  65

ctg caa aga cca gcg ccg tgg ggg cct gcg cta cat ggg aaa gct tgt   294
Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Leu His Gly Lys Ala Cys
    70                  75                  80

ggc atc tgc tcc ctg cag ggc cgt gcc gct gtc ccc acc ttg gct cac   342
Gly Ile Cys Ser Leu Gln Gly Arg Ala Ala Val Pro Thr Leu Ala His
 85                  90                  95                 100
```

```
ctg gcc acc tca cct gca ggt agg aat cct aat acc aac ctg gag gcc   390
Leu Ala Thr Ser Pro Ala Gly Arg Asn Pro Asn Thr Asn Leu Glu Ala
                105                 110                 115

ctg gaa gaa ttt aag aaa ttg gtg cag cgc aag gga ctc tcg gag gag   438
Leu Glu Glu Phe Lys Lys Leu Val Gln Arg Lys Gly Leu Ser Glu Glu
            120                 125                 130

gac att ttc atg ccc ctg cag acg gga agc tgc gtt ctc gaa cac       483
Asp Ile Phe Met Pro Leu Gln Thr Gly Ser Cys Val Leu Glu His
        135                 140                 145

taggcagccc ccgggtctgc acctccagag cccaccctac caccagacac agagcccgga 543

ccacctggac ctaccctcca gccatgaccc ttccctgctc ccacccacct gactccaaat 603

aaag                                                              607
```

<210> 8
<211> 147
<212> PRT
<213> Homo sapiens

<400> 8

```
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
  1               5                  10                  15

Ser Phe Thr Leu Glu Glu Glu Asp Glu Gly Glu Ser Val His Pro Glu
             20                  25                  30

Glu Asn Pro Asp Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu
         35                  40                  45

Trp Gly Gln Glu Ala His Ile Pro Ala Gly Ala Ala Arg Asp Gly Arg
     50                  55                  60

Leu Arg Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Leu His
 65                  70                  75                  80

Gly Lys Ala Cys Gly Ile Cys Ser Leu Gln Gly Arg Ala Ala Val Pro
                85                  90                  95

Thr Leu Ala His Leu Ala Thr Ser Pro Ala Gly Arg Asn Pro Asn Thr
            100                 105                 110

Asn Leu Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln Arg Lys Gly
        115                 120                 125

Leu Ser Glu Glu Asp Ile Phe Met Pro Leu Gln Thr Gly Ser Cys Val
    130                 135                 140

Leu Glu His
145
```

<210> 9
<211> 676
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(552)

<220>
<223> cDNA2098 (676) - conventional form (hOBPIIb-alpha)

<400> 9
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
Met Lys Thr Leu
1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg 102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
5 10 15 20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc 150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
25 30 35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg 198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
40 45 50

aag gtg aca gcc ctg ggc ggt ggg aag ttg gaa gcc acg ttc acc ttc 246
Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr Phe Thr Phe
55 60 65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg 294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
70 75 80

gag gag cct ggc aaa tac agc gcc tat ggg ggc agg aag ctc atg tac 342
Glu Glu Pro Gly Lys Tyr Ser Ala Tyr Gly Gly Arg Lys Leu Met Tyr
85 90 95 100

ctg cag gag ctg ccc agg agg gac cac tac atc ttt tac tgc aaa gac 390
Leu Gln Glu Leu Pro Arg Arg Asp His Tyr Ile Phe Tyr Cys Lys Asp
105 110 115

cag cac cat ggg ggc ctg ctc cac atg gga aag ctt gtg ggt agg aat 438
Gln His His Gly Gly Leu Leu His Met Gly Lys Leu Val Gly Arg Asn
120 125 130

tct gat acc aac cgg gag gcc ctg gaa gaa ttt aag aaa ttg gtg cag 486
Ser Asp Thr Asn Arg Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln
135 140 145

cgc aag gga ctc tcg gag gag gac att ttc acg ccc ctg cag acg gga 534
Arg Lys Gly Leu Ser Glu Glu Asp Ile Phe Thr Pro Leu Gln Thr Gly
150 155 160

```
agc tgc gtt ccc gaa cac taggcagccc ccgggtctgc acctccagag          582
Ser Cys Val Pro Glu His
165                 170

cccaccctac caccagacac agagcccgga ccacctggac ctaccctcca gccatgaccc 642

ttccctgctc ccacccacct gactccaaat aaag                              676


<210> 10
<211> 170
<212> PRT
<213> Homo sapiens

<400> 10
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
  1               5                  10                  15

Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
             20                  25                  30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
         35                  40                  45

Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala
     50                  55                  60

Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65                  70                  75                  80

Met Arg Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Tyr Gly Gly Arg
                 85                  90                  95

Lys Leu Met Tyr Leu Gln Glu Leu Pro Arg Arg Asp His Tyr Ile Phe
            100                 105                 110

Tyr Cys Lys Asp Gln His His Gly Gly Leu Leu His Met Gly Lys Leu
        115                 120                 125

Val Gly Arg Asn Ser Asp Thr Asn Arg Glu Ala Leu Glu Glu Phe Lys
    130                 135                 140

Lys Leu Val Gln Arg Lys Gly Leu Ser Glu Glu Asp Ile Phe Thr Pro
145                 150                 155                 160

Leu Gln Thr Gly Ser Cys Val Pro Glu His
                165                 170


<210> 11
<211> 782
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(537)
```

<220>
<223> cDNA2098 (782) - long form (hOBPIIb-beta)

<400> 11

```
cgc cca gtg acc tgc cga ggt cgg cag cac aga gct ctg gag atg aag   48
Arg Pro Val Thr Cys Arg Gly Arg Gln His Arg Ala Leu Glu Met Lys
  1               5                  10                  15

acc ctg ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc   96
Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe
            20                  25                  30

acc ctg gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg   144
Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met
        35                  40                  45

gtg gtc gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc   192
Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser
    50                  55                  60

cca gtg aag gtg aca gcc ctg ggc ggt ggg aag ttg gaa gcc acg ttc   240
Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr Phe
65                  70                  75                  80

acc ttc atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg   288
Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg
                85                  90                  95

aag acg gag gag cct ggc aaa tac agc gcc tgc ttg tcc gca gtc gag   336
Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Cys Leu Ser Ala Val Glu
            100                 105                 110

atg gac cag atc acg cct gcc ctc tgg gag gcc cta gcc att gac aca   384
Met Asp Gln Ile Thr Pro Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr
        115                 120                 125

ttg agg aag ctg agg att ggg aca agg agg cca agg att aga tgg ggg   432
Leu Arg Lys Leu Arg Ile Gly Thr Arg Arg Pro Arg Ile Arg Trp Gly
    130                 135                 140

cag gaa gct cat gta cct gca gga gct gcc cag gag gga cca cta cat   480
Gln Glu Ala His Val Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His
145                 150                 155                 160

ctt tta ctg caa aga cca gca cca tgg ggg cct gct cca cat ggg aaa   528
Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys
                165                 170                 175

gct tgt ggg taggaattct gataccaacc gggaggccct ggaagaattt           577
Ala Cys Gly

aagaaattgg tgcagcgcaa gggactctcg gaggaggaca ttttcacgcc cctgcagacg 637

ggaagctgcg ttcccgaaca ctaggcagcc cccgggtctg cacctccaga gcccacccta 697

ccaccagaca cagagcccgg accacctgga cctaccctcc agccatgacc cttccctgct 757

cccacccacc tgactccaaa taaag                                       782
```

<210> 12
<211> 179
<212> PRT
<213> Homo sapiens

<400> 12
Arg Pro Val Thr Cys Arg Gly Arg Gln His Arg Ala Leu Glu Met Lys
1 5 10 15

Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe
20 25 30

Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met
35 40 45

Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser
50 55 60

Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr Phe
65 70 75 80

Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg
85 90 95

Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Cys Leu Ser Ala Val Glu
100 105 110

Met Asp Gln Ile Thr Pro Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr
115 120 125

Leu Arg Lys Leu Arg Ile Gly Thr Arg Arg Pro Arg Ile Arg Trp Gly
130 135 140

Gln Glu Ala His Val Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His
145 150 155 160

Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys
165 170 175

Ala Cys Gly

<210> 13
<211> 542
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(297)

<220>
<223> cDNA2098 (542) - short form (OBPIIb-gamma)

<400> 13
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
Met Lys Thr Leu
1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg 102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
5 10 15 20

gag gag gag gat gag gga gga tcg gtg cat cca gaa gaa aat cct gat 150
Glu Glu Glu Asp Glu Gly Gly Ser Val His Pro Glu Glu Asn Pro Asp
25 30 35

gcg gaa gac gga gga gcc tgg caa att cag cgc cta tgg ggg cag gaa 198
Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu Trp Gly Gln Glu
40 45 50

gct cat ata cct gca gga gct gcc cag gag gga cca cta cat ctt tta 246
Ala His Ile Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His Leu Leu
55 60 65

ctg caa aga cca gca cca tgg ggg cct gct cca cat ggg aaa gct tgt 294
Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys Ala Cys
70 75 80

ggg taggaattct gataccaacc gggaggccct ggaagaattt aagaaattgg 347
Gly
85

tgcagcgcaa gggactctcg gaggaggaca ttttcacgcc cctgcagacg ggaagctgcg 407

ttcccgaaca ctaggcagcc cccgggtctg cacctccaga gcccacccta ccaccagaca 467

cagagcccgg accacctgga cctaccctcc agccatgacc cttccctgct cccacccacc 527

tgactccaaa taaag 542

<210> 14
<211> 85
<212> PRT
<213> Homo sapiens

<400> 14
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
1 5 10 15

Ser Phe Thr Leu Glu Glu Glu Asp Glu Gly Gly Ser Val His Pro Glu
20 25 30

Glu Asn Pro Asp Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu
35 40 45

Trp Gly Gln Glu Ala His Ile Pro Ala Gly Ala Ala Gln Glu Gly Pro
50 55 60

Leu His Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His
65 70 75 80

Gly Lys Ala Cys Gly
85

<210> 15
<211> 10664
<212> DNA
<213> Homo sapiens

<220>
<223> hOBPIIa Gene

<220>
<221> modified_base
<222> (8507)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (10635)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (10643)
<223> a, t, c, g, other or unknown

<400> 15
ctctgcttct atattcggcc tccaaatctg tgctcaatgc agcaactgga gtgacccctt 60
aaatacgtaa gtcacagctt gcctttgtca gagctatcca gggtctttca ctcagagcag 120
aagctgaagt cctcgtggtg gtccttaatc cctacatggc cgttccaccc actccccagc 180
ctcatgtgtg gccggtctcc ctggatcatt tgctgtggct gctctgccgt gtgttcccgg 240
aaactgccag cgttctccca cctctgggct ggcactggat gctcctgcca cctggactcc 300
tcttccagct gacgagctca tggcttgctt ccttcatgtc ttaaattcgg tgtttgaatg 360
ccaccttggc gaggctgttc ctcatcaatt catgtaaagg acaaataacc ccttgtctgg 420
ccactcctgt ccaacttgtc cactttgctt tttccatagc actgatcacc atttaaaata 480
atgtatgaac cagtgaccca tcagtgacca gagacagtag attgaacagg tattgcttgc 540
acagtgaaca atcagctcta gaggtttaaa gagaaccaca aagaatggct caggattgca 600
cgagggcagt taaggaagaa ataaacaggg gtggggatcc tttccacaag tgtgggttca 660
gaccgcatgg gagaaggtgt ggttctccca agggaagctg gagaagtttg ctgggttgtc 720
ccagccacag ctggcccacg gtcagggcag aggccagcag gcagggaaat gtaggctggg 780
tctggcaagc aggaagcctc tcttcccccg caaggcaggt gggctggggc tgggaagcct 840
acaagagtca ctgggaaccc acaggtgcag atgccactga atctcaatag gaagccatct 900
gggggggtcc cctgaattca atgtggtgtg tgccgccaga cgtccccaac ttgtgccact 960
gccatttata caggaagaga aagaaaaagg aagaaatgga agcatctgga accagtcatc 1020
ctggacccat gctaggaggt gcttcccccc acgcctcaac caacaaagct tagcattgca 1080
ccaggtgcac aagagaagcg cctgctggct ccagctccat tgcctcggaa ccagccatga 1140
agggtgcgtg tggagctgga ggcaagacat tgatagctgg cactgcaatt cacttattta 1200
ttttgttcat tttaagtccc ctgcacctag aatataagcc ccccaagcac aggacatttg 1260
ttttattgat cgatgtattc cttgtgcccc aaagaatgag aggcatctag aaagtctgca 1320
aaaatcaaac ataaaaatga acctttattc agtcattgtt attttgatgg atatttgagg 1380
catttccaag atttgtaagt aacaattgaa ccccttctcct ggttcatgtg tgggagtgta 1440
tctgttcaaa tgatgcttct gagtggagtt gctgagtctt tggctctagg ttttttttt 1500
ttaagcattt atgctcattg tggtttttaa attaaacatt taaccctgag acactgtaga 1560
ttcccatgca attgtaaaaa gccacacaga gctatcatgt gtatcttcac ctggcttgct 1620
ccagccccaa ccccagcaat gacagacctg ttctccactc ctgcaatctg ctcattgcaa 1680

gaatgtcgtc tattgcaatc ataaaattgt gggattggct ttttttttcc tgtgcagcat 1740
cattctctgg agattcatcc tattgttgca tttatcaata gtttattcca ttttacttct 1800
gagtagtgct ctatggtatg gatgtaccac agtctgttta accattcacc tgttggagga 1860
agtctgtgtt tatagatttg ggctatgaca catgtatagg tttttgcatg gacatcagtt 1920
ttcatttccc tgggacaaag gcccaggggt tctattgctg gattctatgc ttgttacagg 1980
gctcattttg ttttgttttg ctttgtttta ttttgttttt aacctgtcaa gccattttcc 2040
agatccagtt tccatgcatc ctcaccaggc ttcagtatga tcactatgat cttatctcag 2100
ccaccttaat aggtatgtac tgatatatca tggcttttat ttgcatttca ctgatgacta 2160
atggtgttga gcatcttttc atgtgtttat ttgccatctg tatatcctct ctagtcaagt 2220
gtcccttcat gtcttttgtt tacgttctat ttttgaaact gttgagtttt gaaaaattct 2280
ttataaagta tagaaactaa ttctttgttg aatatgtagt ttgtcaatat tttctttcag 2340
tctttggctt gtctttttat tctgttaaca gggtctctta cagagcaaaa ggtttttatt 2400
ttgatgaagt ctattttaac aatttttcct tttatggatc atatttttgg caacaaatct 2460
aaaacctcct gacccagctc catatgtcaa agattttctt gttttctaaa agttttatag 2520
ttttaagttt tatgttgaag tctatgatcc attttgagtt aattttcata cagggtgtga 2580
gaccaaggtt gtggttcttc ttttttcttg gtggttttgt ctgtggatgt ccagttgctc 2640
cagcccattt gctacaaaag ctatctttcc tccactgaat tacttttgca tctttgtaaa 2700
aatgtaattg ggtgtatttg tacaggtctg tttgaggatt ctttattgtc tgtcccattg 2760
atctatgcat ctgtccatgt gctagctata taagtcttga aagggtagcc tgtagctggg 2820
tgtggtggcg agagcctgta atcccagcta ttcgggaagt ggaggcagga gaatcgattg 2880
aacccaggag gtggaggttg cagtgagtca agatcgtgcc actgcactcc agcctgggtg 2940
acagagtgaa ctccatataa aaaaataaaa acaaaaataa agtagcgtga tttctcccac 3000
cttattcttt tttaaaaaag ttttagctat tccagttcct ttgcctttcc atataaattt 3060
tagaataatc ttgtctatat ctaccacaat tctttctgga attttgatag aaatttttgtt 3120
acatctttat attatttgag agaaatgata tttttactat gttgagtctt ctaatccatg 3180
gatataacgt ctctccattt gtttagatct tctttgattt attttataat cattgcattg 3240
ttttcagcat acaaatccag catatgggtt gttagactta tgcctaggca tctcattttt 3300
ttagccatta taaatagtac tgtgttttta agtttagggt ccattactag tatgtagaca 3360
cacaattgat ctttgtatat ttatcttgtg tcctgcaacc ttgctgaact cacttactag 3420
ttctaggagg tgtattgttt tctttttgttt tgtttttcaa tttcttggga ttttctacag 3480
agataataat gtcatctgca gatgcagttt tcttccttcc tttccaattt gtatgtcttt 3540
aatttccttt ttaaaaaacc tatattactc tgactagaac tttctgtact atgttaaata 3600
caagtggtag agtggacatc cttgccttgt ccctgatgtt aaagagaaag catttgtaac 3660
tgagtatccc agcctcaaaa tgtgcttaaa aacttttttc ctttcttgct ttcagccttg 3720
aaacatactt cgaaactctt tatttctccc tttcccacca ggcacttccg tgagcagtgc 3780
tcgcttatct aattatgtgc ttacttagaa attccagggg ccaattttga aacaaaccag 3840
gcagagagac ccagctgcag aatcctccct cttaggggga gttacaggta gcctaccact 3900
tcccggctga aatcaggatg acgcaaacca gacctccgga cagacgattg atgactcaca 3960
ataaccatca gaacaagatg cagaccaaca tcctcctgca ccattcccac atatttccca 4020
caccttttcc tccttaaacc ccttcgctca gtccagaaaa tctgaatggt cttttaaagg 4080
catgggtctg gccattcccc aactgccagt atttgaataa agctgctttc cttttaccac 4140
acctcacttc tcatgccttg acttctgagc agcgagcagc tggacttgag ccagttacac 4200
atcagtcttt caccattaga aataatgtag ccataggttt ttttttcgta gatgttcttt 4260
ttcaagttaa agaagttctc ttctattcct atttttctga gaggtgttat cccgaatgag 4320
tgttgaattt tgttaaatac ttttaacaac ccaacaggaa ccaccatcag aagccatcca 4380
gagaaacgca ccaggccaca aagcactggg ggccagggat cttgcccctg ctgtctgcca 4440
tgggttgacc cccagcctcc aaccctacca tcccctgacg gtgtctgcag cagttgaacc 4500
caaccagcat ctaaaagaac acagttggtg aacgagactg ggacacaggg caagatgggt 4560
ggacaatggg aggctcctgg agagcacccg taccagcgag catagaattc atgggggtga 4620
cctgttccct gaagcatctg cgcgtgttgt tccagcattt tcttcaagga ttgagccagc 4680
agcaccagtg tcatacggtg cttaaatcaa tgattcacag ccaaccaatg aaatacaagg 4740
tgccggctgg gcgcggtggc tcacgcctgt aatcccagca ctttgggagg ccgaggcggg 4800
cggatcacga ggtcaggaga tcgagaccat cccggctaaa acggtgaaac cccgtctcta 4860
ctaaaaatac aaaaaaaatt agccgggcgt agtggcgggc gcctgtagtc ccagctactt 4920
gggaggctga ggcaggagaa tggcgtgaac ccgggaggcg gagcttgcag tgagccaaga 4980
tcccgccact gcactccagc ctgggcgaca gagcgagact ccgtctcaaa aaaaaaaaaa 5040
aaaaaaaaaa aaaagaaata caaggtgcct ggggtacagg caacaaaaat ggggaacgga 5100
gaaatcttcc cgttgaggta gagactccag cctgtttgct catctctggt cctgaagagc 5160

ccagtccccg ccctaaggat ggggtttctg ctcggcagca cttgccgtga gaggggtgag 5220
gcactgggtc acgccagccc tttctttata gcccccaggg tttattagag tgtgcattag 5280
tattatttag caagcacttg agggtgtctg atcttgggcc agagaggcac aaagattgtg 5340
tgggcccagc acctgcctgc agagcgtggg tcagccttgg gtcccagggc agatgacgca 5400
ggcccgggaa ctacagggtc caggagggag aaggcaaagt tctggctcag gttggctggg 5460
gatgaggcca gcggagccag gtgcccaagg gagctcagcc acaaactctg agcacaggct 5520
ggcaggtggc tcttgatgct catgccaccc atttatctaa agggatgaga ttcaaggcct 5580
gtcctggtgg ctgggggccg tcgaagctga cgagagaggg gatgtagagt gaatgtatat 5640
tccactctac cacttgtatt taacagggag atggatgatg aacacgtgca ggaggaaaca 5700
ggcaggacaa tccagagaga tcacgtgttc tgaggacagc acagccaggc tccggtacgg 5760
agtgaagcgg ggtggggcag gcggcggggt ccctcatatg gcccgaggag gccgtatata 5820
tactgacctt gagccacaca atagtgccct tctctgcccc taggaagctc gagtgcaggg 5880
tccaggtggg gaaaatcaat gcagagtggg tcccagagtg ggcggaagct tgggctctag 5940
ggcgtgcggg actcagctgg cagcagcccc acatctctat ggttctaaag cccagtccat 6000
ttctgctcag cagggagatg gccagttccc cagaggactt gcccagggtc ccagccgtgg 6060
ccctgggagg gctcaggggt gagggaggga gaattccgag ccgtgggccc tccctctgtg 6120
gctcggggtg cagacgtgca ctaccctgcc ctgtcctctg gagtcctccg ggctcatcca 6180
agtgggcgca tttgggtgca aaccctggac agcatcgtgt gtcttttctt ctctggctgg 6240
cactgaattg ctgttcacca gatgccagca atgaggacac tccccctccc ggatgcagga 6300
ctggttgcca ccatggggag gggtgcaccg tgccacgtgc tcccaggaca ctggccaggg 6360
ggctataaag aacatctcga gaggagccag cacagccttg ttcagacgcc cagtgacctg 6420
ccgaggtcgg cagcacagag ctctggagat gaagaccctg ttcctgggtg tcacgctcgg 6480
cctggccgct gccctgtcct tcaccctgga ggaggaggat gtgagctggg ttggcgtggg 6540
cggatggagg agccaggtgg actcctgggc agggggcagt gccaggggcc ctgctttagg 6600
aggtgtcact taagcctggg gtggtggtgg aggggtccta ttgttccttc agcagacaat 6660
gctccccatg aggcccaggg ccggagcagg ctcggctcag gggttcctgc tgcactgacg 6720
cctgaagccc gaaggtctcg cagggttggg ccctgtggag ggagggctca cctggtgctg 6780
gggcccgggg gtccatgggg tgcagacatg ccctccttcc actgggggct gggagccctg 6840
agcagggggc tggctctaac tcactccagc tgagctctaa ctaaggtgca ggaacccagc 6900
ctgcctttag gggtggcagc cgggcaccat gggtgtctgg ttatagctgc aggcctgagt 6960
gccagggtca gagtagaatc tgggccaccc atggtgggct cacggccttg gcctgctcca 7020
gatcacaggg acctggtacg tgaaggccat ggtggtcgat aaggactttc cggaggacag 7080
gaggcccagg aaggtgtccc cagtgaaggt gacagccctg ggcggtggga acttggaagc 7140
cacgttcacc ttcatgtgag tgttgcccac tgcagggccc ctcaggccac tttcgctccc 7200
cgccccagac ccacctggtg cccattgccc catccacatt tcgggtgttg ggaagagtca 7260
ccccctgcct tggagggaaa cagccagggc atcctgaagc tcggtggggt ggggggggcag 7320
tggaattttc aggttgccgg gtcagggcca tgcaccaggt gagctgagga tgggccaggt 7380
gtgtcctggg agccgctgcc cgcgtgtctc ctgttttcca ggagggagga tcggtgcatc 7440
cagaagaaaa tcctgatgcg gaagacggag gagcctggca aatacagcgc ctgtgagccc 7500
ctccccgacc ccccactccc catgcccaac cccggatgca ccagccccac tgcaggtgga 7560
gagtgcccag gccacacttc tgccagggtc ccagccctgc ccacctccaa ggaggggctg 7620
gcctctcctt cctggggggc tggtggccct gacatcagac accaggtgtg acaggcttgt 7680
ccacagtaga gatggaccag atcaagcctg ccctctggga ggccctagcc attgacacat 7740
tgaggaatcc gagtgttagg gaccaggagg ccgagggtta gggatgggaa gccaaggctg 7800
agggtttggg atcaggaatc cgagggttag ggacagggaa gtggggcagg agcagctgct 7860
ggagctggga aggccggact ctagtcctgg acgtgctctg gccttgtggc tccattactt 7920
gcattgggac cttccgagag gaggctcctg cctccgtgtc cgggtccatg ctgtgcggag 7980
cagccaggcc tggctcaggc tgtccagggc acctgggtga ccactgaaac attcctgagt 8040
gtttcttcgt gtggtcctga gtgctctctc cgggaatgag ggcactgaag acccatcttc 8100
tctgtcatct acagatgggg gcaggaagct catatacctg caggagctgc ccgggacgga 8160
cgactacgtc ttttactgca aagaccagcg ccgtgggggc ctgcgctaca tgggaaagct 8220
tgtgggtgag gggcccgctg gggcctgcat gtcctgcccc atggtctctg cctccagaag 8280
ccagtggaac caccatcatc acgccctggc acgggggggaa aaggaagccc cctgcgccgg 8340
ccttcgtgtg ctaggcacca agcgctgccc tggatggctg gtccaagttc ctgaagtggg 8400
agtggggtgg gccaggcagg gacagacacg gccctcggtg acgtgaacct gccaagggcc 8460
gcttgtgggg tctcaggtgt aggggcctca ccttaagggg gaggtancat cttaacagag 8520
ctcttcatgg ggcagggact ctccagggcg gcagggcagc cagtgcctct gggacacaag 8580
gtccctccag gtgagggttg tgaccctgca gagtggcttt gggagctgcc caggtccccc 8640

tggggttgct gagtggcttg gaccctgcca ctgtcccctt tcctggggac ctctcacctg 8700
ggcggtggcc gtctcctctg tccccagtcc caccctgag ctcttgtcca ttctcaggcc 8760
tcctctcccc cttgcctggt gctggacagt tgccatctct tctgtcccca gccccacccc 8820
tgagctctga tccactctcg ggcctctccc ccgtcctgat gctgggccgt ggtcgtctcc 8880
ttttagcatc tgctccctgc agggccgtgc cgctgtcccc acgtcggctc acctggccac 8940
ctcacctgca ggtaggaatc ctaataccaa cctggaggcc ctggaagaat ttaagaaatt 9000
ggtgcagcac aagggactct cggaggagga cattttcacg cccctgcaga cgggtgagga 9060
cggctgtgcc cagtaccccg tgttcccctg tgtctctgtg tgatctccag tgtcccatga 9120
ccctcgtgtc ctcccatgtc ccccgcattc cccatgtgcc ccgagtctcc tcgcaggggc 9180
tctgggccct gcttagcatc ctcgtcgttg gagggtctgc actctgggct gcgatggggt 9240
ctggggctcc gcgctctggg ctgcgatgcg gtctggggct ccgcgctctg ggctgcgatg 9300
gggtctgggg ctccgcgctc tgggctgcga tggggtctgg ggctccgcgc tctgggctgc 9360
gatggggtct ggggctccgc gctctgggct gcgatggggt ctggggctcc gcgctctggg 9420
ctgcgatggg gtctggggct ccgcgctctg ggctgcgatg gggtctgggg ctccgcgctc 9480
tgggctgcga tggggtctgg ggctccgcgc tctgggctgc gatggggtct ggggctctga 9540
gctctgggct gcgatggggt ctgggccctg gtctagggcg ccttctaatc cctgggtttt 9600
tcttggtctc tgcaggaagc tgcgttctcg aacactaggg tgagtgagcc tttaggaggg 9660
cactggacaa gcccagagtc ctgggttccc ggggttcgag ggtacatctg ctctggccct 9720
tcccatccca cacagccagg gagacccccc cagggtcagg cacgaggttg gcacctcaga 9780
gtctgcccac ccaaaattcc tgggacattc gggaagtcct ttgttttacc attcctgcac 9840
ctgccagccc gagtgagggt cctcctcggc ctttccacag cgaggcctcc ctccggctcc 9900
ctcaggtgtc agctcggccc atcgcccctt gcacctgtcc ccactcggtc tactcccccc 9960
cacccactca ccaaggatgc tcagaggcct gcccagtcat tggaggagct gaggctgctc 10020
tggagccccg aggctgccca gcagtggccg atgtggaagc tgcagagcct ggggagggag 10080
ctctgggcct ggcctctgcc ctaccctgca gcctccctga cctctgctcc tcttcccagc 10140
acagcagccc ccgggtctgc acctccagag cccaccctac caccagacac agagcccgga 10200
ccacctggac ctaccctcca gccatgaccc ttccctgctc ccacccacct gactccaaat 10260
aaagagcttc tcccccagct ctgggcaggc ctatctgtgg ggacggaggg gctcgcaccg 10320
gctccctggg aggcctgctg ggaggggggag ccacaaggga agctggggag atgttggacc 10380
tagcaagggc caaaccacaa gaggcacgat gtccaacagg ctgtggggct gcgtgatgtc 10440
ctggaggggc ctccggaaga gccgccctcg gatctcaggc tcagcttggg acgggcaggg 10500
ctctgggcag gagcacttgg gaagccactg ggagggccgg agtggggaca cggcgtcaga 10560
cctgcatcag tgggtccacc acagggctcc cgccccgggg gtctcagtgg gatcctccga 10620
gcgggtccac tttancatcc tgngctcagt ggcatcactg gata 10664

<210> 16
<211> 11851
<212> DNA
<213> Homo sapiens

<220>
<223> hOBPIIb Gene

<400> 16
tctctgcttc tatattcggc ctccaaatct gtgctcaatg cagcaactgg agtgacccct 60
taaatacgta agtcacagct tgcctttgtc agagctctcc agggtctttc actcagagca 120
gaagctgaag tcctcgtggt ggtccttaat ccctacatgg cagttccacc cactccccag 180
cctcatgtgt ggccggtctc cctggatcat ttgctgtggc tgctctgctg tgtgttcccg 240
gaaactgcca gcattctccc acctccaggc tggcactgga tgctcctgcc acctggactc 300
ctcttccagc tgacgagctc atggcttgct tccttcatgt cttaaattcg gtgtttgaat 360
gccaccttgg cggggctgtt cctcatcaat tcatttaaag gacaaataac cccttgtcct 420
gacactcctg tccaacttgt ccactttgct ttttccatag cactgatcac catttaaaat 480
aacgtatgaa ccagtgaccc atcaggatcc agagacaata gattgaacag ggattgcttg 540
cacaatgaac aatcagctct agaggtttaa agagaaccac aaagaatggc tcaggattgc 600
acgagggcag ttaaggaaga aataaacagg ggtggggatc ctttccacaa gtctgggttc 660
agacctcatg ggagaaggtg tggttctccc aagggaagct ggagaagttt gctgggttgt 720
cccagccaca gctggcccac ggtcagggca gaggccagca ggcagggaaa tgtaggctgg 780

gtctggcaag caggaagcct ctctccccac ccaaggcagg tgggctgggg ctgggaagcc 840
tgcaagagtc actgagaacc cacaggtgca gatgccactg aatctcaata ggaagccatc 900
tgggggcggt cccctgaatt caatgtggtg tgtgccgcca gacgtccccg acttgtgcca 960
ctgccatttg cataggaaga gaaagaaaaa ggaagaaatg gaagcatctg gaaccagtca 1020
tcctggaccc atgctaggag gcgcttctcc ctacgcctca accaacaaag cttagcattg 1080
caccaggtgc acaagagaag cgcctgcttg ctccagctcc attgcctcgg aaccggccat 1140
gaagggtacg tgtggagctg gaggcaagac attgatagct ggcactgcaa ttcacttatt 1200
tattgtgttc attttaagtc ccctgcacct agaatataag ccccccgagc acaggacatt 1260
tgtttcattg atcgatgtat tccttgtgcc ccaaagaatg agaggcatct agaaagtctg 1320
caaaaatcaa acataaaaat gaacctttat tcagtcattg ttattttgat ggaaatttga 1380
ggcatttcca agatttgtaa gtaacaattg aacccttctg ctggttcatg tgtgggagtg 1440
tatctgttga aatgatgctt ctgagtggag ttgctgagtc tttggctcta ggtttttttt 1500
ttttaagcat ttatgcttat tgtggttttt aaattaaaca tttaaccctg agacattgta 1560
gattcccatg cagttgtaaa aagccacaca gagctatcat gtgtatcttc acctggcttg 1620
ctccagcccc aaccccagca atgacagacc tgttctccac tcctgaaatc tgctcattgc 1680
aagaatgtcg tctattgcaa tcataaaatt gtgggattgg ctttttttt tcctgtgcag 1740
catcattctc tggagattca tcccattgtt gcatttatca atagtttatt ccattttact 1800
tctgagtagt gctctatggt atggatgtac cacagtctgt ttaactattc acctgttgga 1860
ggatgtctgt gtttatagat ttgggctatg acacatgtac aggtttttgc atggacatca 1920
gttttcattt ttctgggaca aaggcccagg ggttctattg cggggttcta tgcttgttgc 1980
agggtttttt ttttttaaacc tgtcaagcca ttttccagaa actgtccagt ttccatgcat 2040
cctcaccagg cttcagtatg atcactatga tcttatctca gccaccttaa taggtatgta 2100
ctgatatatc atggctttta tttgcatttc actgatgact aatggtgttg agcatctttt 2160
catgtgttta tttgccatct gtatatcctc tctagtcaag tgtcccttca tgtcttttgt 2220
ttacattcta tttttgaaac tgttgagttt tgaaaaattc tttataaagt atagaaacta 2280
attctttgtt gaatatgtag tttgtcaata ttttctttca gtctttggct tgtcttttta 2340
ttctgttaac agggtctctt acagagcaaa aggtttttat tttgatgaag tctattttaa 2400
caatttttcc ttttatggat catatttttg gtgacacatc taaaatctcc tgacccagct 2460
ccatatccca aagattttct ttctgttttc taaagctttt atagttttaa gttttatgtt 2520
taagtctatg atccattttg agttaatttt cttgcagggt gtgagaccaa ggttgtggtt 2580
cttctttttg tttggtggtt ttgtctgtgg atgtccagtt aatccagccc atttgctaca 2640
aaagctatct tccattgaat tgcttttgca tctttgtaaa aacttaactg ggtatatgcg 2700
tgcaggtctg tttgaggatt ctttattgtc tgtcccattg atctatgcat ctgtccatct 2760
gctagctata taatgagtct tgaaagggta gcctgtagct gggtatggtg gtgtgcatct 2820
gtaatcccag ctacttggga ggcggaggca ggagaatcac ttgaatctgg gaggtggagg 2880
ttgcaatgag tcgagattgt gctactgcac tccagcctgg gtgacagagc aagactctgt 2940
ctctaaaaaa aagaaagaag aagaaaagga aagaaagggt agcctgattc ctcccacttt 3000
attcttaaaa aaaaaagttt tagttattct agttcctttg cctttccata taaatgttag 3060
aataatcttg tctatgtcta caaaaattcc ttctggaatt ttgatagaaa ttgtgttaaa 3120
tctttatatt atttgagaga aatgacgctt ttattatgtt gagtctccta atccatggat 3180
atagcacatc tctccatttg tttagatctt ctttgattta ttttataatc attgcattat 3240
tttcagcata caaatccagc atatgggttg ttaaacttat gcctaggtat ctctttttt 3300
tagccactat aagtagtatt gtgtttttaa gtttagggtc cgttactagt atgtagacac 3360
acaattgatc tttgtatatt tatcttgtat cctgcaacct tgctgaactc gcttaatagt 3420
tctaggaggt gtattgtttt cttttgtttt gtttttcagt ttcttgggat tttctacaga 3480
gacaatcatg tcatctgcag atgcagacag ttttcttcct tcctttccaa tttgtatgtc 3540
tttaatttcc tttttaaaaa acctttattg ctctgactag aactttctgt actatgttaa 3600
atacaagtgg tgaaagtgga catccttgcc ttgtccctga tgttaaggag aaagcgtttg 3660
taactgagta tcccagcctc aaaatgtgct taaaaacttt tttcctttct tgctttcagc 3720
cttgaaacat acttcgaaac tctttatttc tccctttccc accaggcact tctgtgagca 3780
gtgctcgctt atctaattat gtgcttactt agaaattcca ggggccaatt ttgaaacaaa 3840
ccaggcagag agacccagct gcagaatcct ccctcttagg gggagttaca ggtagcctac 3900
cacttcccgg ctgaaatcag gatgactcaa accagacctc tggacagacg attaatgact 3960
catgataacc attggaacaa gatgcagacc aacatcctcc tgcaccattc ccacatattt 4020
cccacacctt ttcctcctta aacccctttcg ctcagtccag aaagtttgaa tggtctttta 4080
aaggcatggg gcctggccat tcctctactg ctagcatttg aataatgctg ctttcctttc 4140
accacacttc acttctcatg ccttgacttc tgagcagcga gcagctggac ttgagccagt 4200
tacacatcag tctttcacca tgagaaatga tgttagccat aggtttttg tagatgctct 4260

ttgtcaagtt aaggaggttc tcttctattc ctacttttct gagagttgtt ttcctgaatg 4320
cgtgatgaat ttagtcaagt acttttttctg cattgatcga tatgatcatg taatttttct 4380
tcttaacata ttaatatggt tgatttttga gtattgtacc aggtttccat ccctggaata 4440
aactgctttt ggtcatggtg tagaattaat tttatatatt attatttgct aatattttgt 4500
aaaggatttc tgtatctata ttcatgaggt atattgttct gcagctttcc tttttgtgtt 4560
gtcaggtttt ggttgggaga tattctctca tcttctgaga ttatgcagag ttggtgttat 4620
ttatttttta aatggttggt agaattttct aatgaaacca tatggacatg aagaattatt 4680
tttggaagct tttaaaatta aaatgttaac tttaattgtc atagagctat taaagttatc 4740
tatcttatat tggcattgtg ttttttgaga aattggttca tttcacctat gttgtcaaat 4800
atatgtggga aggtctcttc gtagtattcc cttattatcc tttgatgtct gcagggtctg 4860
tagtgatagt ctctgttttg ttccagatgt tggttatttg tgtctttctt tctttttttg 4920
tttgtcagtc tttctagaac tgcctcaatt ttattaattt ttccagagaa ctaacccttg 4980
tttcattgtt tttctctgtt gtttttctgc tttcgatttt ttctgctttc tgcttttgat 5040
ttattaattg ctgctcttat ctttgttatt tcttttcttc ttgctttggg tttatttgc 5100
tcttcttttt ctaggttctt gaagtgggag cttagatagt tgattttgga cttctctttt 5160
ctaatatatg catttaatgc ctctcagcac tgatttagct gtgtctcaca aattttgata 5220
tgtttttgtt tgtttgtttt tgaggtagag tctcgctctg tatcccaggc tggagtgcag 5280
tggcactatc ttggctcact gcaacctcca cctcctgggt tgaagtaatt ctcctgcttc 5340
agccccctaa gtagctggga ttacaggtat acgccaccat gcccagcaat ttatttattt 5400
atttatttt gtatttttag tagagatgag gttttgtcat gctgtccagg ctggtctcga 5460
actcctgacc tcaagtgatc tgcccacgtg gggctccaaa gcactgagat tataggcata 5520
agccactgtg cccagcctgt ttttgttttt tttgagatgg agtcttgctc tgtcacccag 5580
gctggagtgc agtggcatga tagcccactg caacctccgc ctcccgggtt caagtgattc 5640
tcctgcctca gcctcccgat tagctgggat tacaggcgtg caccaccaca ctcagctagt 5700
ttttgcattt ttggtagaga tgggtttttcg ccatgttggc caggctggtc tcgaattcct 5760
gatcgcaagt gatccacccg cctgggcctc ccaaagtgct ggaattatag gcacaagcca 5820
tcacgcccag ctgatatgtt gtattttcat tttcatccag cctagtatat ttttaaattt 5880
accttcaggc ttcctctttg acctgtggat tattcacagg tgtgttgttt ccaagcatct 5940
gaagaaagat tttccattat ctttctggca ttgatttata gttagattcc atgtggtcag 6000
agactaccct ctatatcatt taaatttttta aaaatttctt gaggctcatt ttctcatcca 6060
ggatatggtc tatcttggta tatattctgg tgcacttgaa aagaatgtgt gtagtgctgt 6120
tgccaggtgg tgtgttctac aaatgtcaat tcgattatgt taattgatgg gtggccgagt 6180
tctttgatgt ccctgctggt tttctgtcta gttgctctga gagaagggta ttgaagtctc 6240
caactataat tgtggctttt taaatttctc ctttcagttc tgttttgctt cacatattgt 6300
gcagctaatg tttggtgcac acacgtttag gattactatg tcttctttgc agaatggccc 6360
acttattatt gtataatttc cctctctgtt aattagtctt tgttctaaag tctattttac 6420
ctgttattga aataatgctt ctgctttctt ttgattaatg tttatatatc ttttttcatg 6480
tttttacttt caacctgtct ctattgttac atttgaagta agtttcttgt agacggcata 6540
tagttgtata atggctttta attcactttg ccaatcactg tcatttaata tttaaaccat 6600
ttacatttaa tataattatt aacatgctag agtttgtatt atttttatag ggttgctata 6660
acaaaatact accaactggg tggcttagaa caaaaattta ttttctcaca ttctggtggc 6720
cagaagtctg agatcaagtg agatttggaa ggtagaaaaa ccaacaaaat atgttaatga 6780
tgttaccatt gtgggcaact gggactcagt ccttctgagg gccatctgag aaatagggtg 6840
gaaccatatt tgattgttcc acttagaagg aagagtccag accatttatc ccatcagtta 6900
ctatccctac agactgagga tgctgattca cttgcacatc tgggtttcat cagtgtcagg 6960
ggaagcacag aagaaaaggt gcaggcactt aaggtgggaa gctgtgaata cgtccgtgca 7020
gggctctagg gcgtgcggga ctcagctggc agcagcccca cgtctctctg gttctaaagc 7080
ccagtccatt tctgctcagc agggagatgg ccagttcccc agaggacttg cccagggtcc 7140
cagccgtggc cctgggaggg tgcaggggtt ggtgagggag tatcccgagg ctgtgggccc 7200
tgcctctgtg gcttggggtg cagatggaca ccaccctgcc ctgtcctctc gagtcctctg 7260
ggctcatctg agtgggcgcg ttcgggtggt gcaaactctg gatggcatcg tgtgtctttt 7320
cttctctggc tggcactgac ttgctgttca ccagatgcca gcaatgagga cactcccct 7380
cccggatgca ggactggttg ccaccatggg gaggggtgca ccgtgccacg tgctcccagg 7440
acactggcca gggggctata aagaacatct cgagaggagc cagcacagcc ttgttcagac 7500
gcccagtgac ctgccgaggt cggcagcaca gagctctgga gatgaagacc ctgttcctgg 7560
gtgtcacgct cggcctggcc gctgccctgt ccttcaccct ggaggaggag gatgtgagct 7620
gggttggcgt gggcggatgg aggagccagg cggactcctg ggcaggggggc agtgccaggg 7680
gccctgcttt aggaggtgtc actgaaggct ggcttctgac cccgtgctcc cagcctgggg 7740

tggtggtgga ggggtcctat tgttcctcca gcagacacgg ccccccgagg cccagggccg 7800
gagcaggctc gtctcagggg ttcctgctgc actgacgcct gaagcccgaa ggtctcgcag 7860
ggttgggccc tgtggaggga gggctcacct ggtgctgggg cccgggggtc catggggtgc 7920
agacatgccc tccttccact gggggctggg agccctgagc agggggctgg ctctaactca 7980
ctccagctga gctctaacta aggtgcagga acacagcctg cctttaaggg cagcagccgg 8040
gcaccatggg tgtctggtta tagctgcagg cctgagtgcc agggtcagag tagaatctgg 8100
gccacccatg gtgggctcac ggccttggcc tgctccagat cacagggacc tggtacgtga 8160
aggccatggt ggtcgataag gactttccgg aggacaggag gcccaggaag gtgtccccag 8220
tgaaggtgac agccctgggc ggtgggaagt tggaagccac gttcaccttc atgtgagtgt 8280
tgcccactgc agggcccctc aggccacttt cgctcccctc cccagaccca cctggtgccc 8340
attgccccat ccacgtttcg ggtgttggga agagtcaccc cctgccttgg agggaaacag 8400
cctgggcatc ctgaagctcg gtggggtggg ggacagtgga attttcaggt tgccgggtca 8460
gggccatgca ccaggtgagc tgaggatggg ccatgggtgt cctgggagcc gctgcccgcg 8520
tgtctcctgt tttccaggag ggaggatcgg tgcatccaga agaaaatcct gatgcggaag 8580
acggaggagc ctggcaaata cagcgcctgt gagcccctcc cccactccca cccccaccct 8640
cccccaccgc caaccccagt gcaccagcct ccacaggtag agagtgccca ggctgccctt 8700
ttgccagggc cccagctctg cccacctcca aggaggggct ggcctctcct tcctgggggg 8760
ctggtggccc tgacatcaga caccgggtgt gacaggcttg tccgcagtcg agatggacca 8820
gatcacgcct gccctctggg aggccctagc cattgacaca ttgaggaagc tgaggattgg 8880
gacaaggagg ccaaggatta ggtacgggga ggctgagggt tagggatggg gaagctgagg 8940
gttaaggatg gggaagctga gggttaagga tggggaggct gagggttagg gacgggggcc 9000
aagggttagg gatcgggaag ctgagggtta gggatgagga ggccgagggt tagggatggg 9060
gaggccgagg gttagtgctg cggaagctga ggattagaga tggggaggct gaggggggaag 9120
atgagggtta gggacaggga agtggggcag gagcagctgc tggagctggg aaggccggac 9180
tctagtcctg gacgtgctct ggccttgtgg ctccattact tgcattggga ccttcctgag 9240
aggaggctcc tgcctccgtg tccgggtcca cactgtgcgg agcagccagg cctggctcag 9300
gctgtccagg gcacctgggt gaccgctgaa acattcccga gtgtttcttc atgtggcccc 9360
gagtgttctc tctgggaatg acccattttc tctgtcatct acagatgggg gcaggaagct 9420
catgtacctg caggagctgc ccaggaggga ccactacatc ttttactgca aagaccagca 9480
ccatgggggc ctgctccaca tgggaaagct tgtgggtgag gggcttgctg gggtctgcgt 9540
gtcctgcccc acggtctctg cctctggaag ccggtggaac cagcaccacc atgccctggc 9600
atgggggaaa aggaagcccc ctgtgccggc tttcatgtgc cgggcaccgc tgggcagccg 9660
gtccaagtac ctgaagtggg aatgggatgg gccaggcagg gacagacatg gccctcagtg 9720
acatgaacct gccaagggcc acttctgggg tctcgggtgt aggagggtca ccttaagggg 9780
gagggtcacc ttaagcgtcc taagagagct cttcatgggg cagggaccct ccagggcagg 9840
agggtggccg gtgcctctgg gagacaaggt ccctccaggt gagggctgtg accctgcagg 9900
gtggcattgg gagctgccca ggtcccctgg ggttgccgag tggcttggag agtccccgcc 9960
actgtccccc ttcctgggga cctctcatct gggcagtgac tgtctcctct gtccccagtc 10020
ccacccctga gctctcatcc attctcaagt ctcctctcct gcttgtccgg tgctgggcta 10080
tggccatctc ttctgtcccc ggccccaccc cagagctctg gtccactctt gggtctctcc 10140
ccttgtcctg gtgctgggca gtggccatct cctctgtccc cagccccacc cccgagctct 10200
gatccactct caggcctctc cccctgtcct ggtgctgggc cgtggtcgtc tcctttcggc 10260
atctgctccc tgcagggccg tgccgctgtc cccacgtcgg ctcacctggc cacctcacct 10320
gcaggtagga attctgatac caaccgggag gccctggaag aatttaagaa attggtgcag 10380
cgcaagggac tctcggagga ggacattttc acgcccctgc agacgggtga ggatggctgt 10440
gcccagtccc ctgtgtccct ctgctgtgtc tgtctgctat ctccagtgtc ccatgacccc 10500
catgtcctcc catgtccccc gcattcccca tgtgccccga gtctcctcgc aggggctccc 10560
gggccctgtt tagcgtcctc ctcattggag gctctgtgct ctgggctgcg atggggtctg 10620
gggctccgcg ctctgggctg cgatggggtc tggggctccg cactctgggc tgcgatgggg 10680
tctggggctc cgcgctctgg gctgcgatgg gctctggggc tctgagctct gggctgtgat 10740
ggggtctggg ccctggtcta gggcgccttc taatccctgg gtttttcttg gtctctgcag 10800
gaagctgcgt tcccgaacac tagggtgagt gagcctttag gagggcactg gacaagccca 10860
gagtcctggg ttcccggggt tcgagggtac atctgctctg gcccctccca tcccacacag 10920
ccagggagac ccccccaggg tcaggcacga ggttggcacc tcagagtctg cccacccaca 10980
gttcctggga cattcgggaa gtcctttgtt ttatcattcc tgcacctgcc agcccgagtg 11040
agggtcctcc tcggcctttc cacagcgagg cctccctccg gctccctcag gtgtcagctc 11100
agcccatcgc cccctgcacc tgtccccacg cggtccactc cccaccatcc acccaccaag 11160
gatgctcaga ggcctgccca gtcattggag gagctgaggt cgctctggag ccccgaggtt 11220

gcccagcagt ggccgatgtg ggggctgcag agcctgggga gggagctctg gccctggcct 11280
ctgccctacc ctgcagcctc cctgacctct gctcctcttc ccagcacagc agcccccggg 11340
tctgcacctc cagagcccac cctaccacca gacacagagc ccggaccacc tggacctacc 11400
ctccagccat gacccttccc tgctcccacc cacctgactc caaataaagt ccttctcccc 11460
cagctctggg caggcctatc tgtggggaca gaggggcttg cacccgctcc ctgggaggtc 11520
tgctgggagg gggagccaca agggaagctg gggagatgtt ggacctagca agggccaaac 11580
cacaagaggc acgatgtcca acaggcctga ggggctgcgt gatgtcctgg aggggcctcc 11640
ggaagagccg ccctcaggtc tcaggttcag cttaggacag gcagggccct gggcaggagc 11700
gtttggaaag ccactgggga ggacaggagc ggggacacgg cgtcagggct gcagcagtgg 11760
ggccaccgca gggctcccgc ctcaggggtc tcagagggat cctccaagct cccccatttt 11820
agcagcctgg gctcagtggc agcactggag a 11851

<210> 17
<211> 253
<212> DNA
<213> Homo sapiens

<400> 17
cactcccccc cccggatgca ggactggttg ccaccatggg gaggggtgca ccgtgccacg 60
tgctcccagg acactggcca gggggctata aagaacatct cgagaggagc cagcacagcc 120
ttgttcagac gcccagtgac ctgccgaggt cggcagcaca gagctctgga gatgaagacc 180
ctgttcctgg gtgtcacgct cggcctggcc gctgccctgt ccttcaccct ggaggaggag 240
gatgtgagct ggg 253

<210> 18
<211> 154
<212> DNA
<213> Homo sapiens

<400> 18
cctgctccag atcacaggga cctggtacgt gaaggccatg gtggtcgata aggactttcc 60
ggaggacagg aggcccagga aggtgtcccc agtgaaggtg acagccctgg gcggtgggaa 120
cttggaagcc acgttcacct tcatgtgagt gttg 154

<210> 19
<211> 91
<212> DNA
<213> Homo sapiens

<400> 19
tgttttccag gagggaggat cggtgcatcc agaagaaaat cctgatgcgg aagacggagg 60
agcctggcaa atacagcgcc tgtgagcccc t 91

<210> 20
<211> 126
<212> DNA
<213> Homo sapiens

<400> 20
ggtgtgacag gcttgtccgc agtcgagatg gaccagatca cgcctgccct ctgggaggcc 60
ctagccattg acacattgag gaagctgagg attgggacaa ggaggccaag gattaggtac 120
ggggag 126

<210> 21
<211> 131
<212> DNA
<213> Homo sapiens

<400> 21
tcatctacag atgggggcag gaagctcata tacctgcagg agctgcccgg gacggacgac 60
tacgtctttt actgcaaaga ccagcgccgt gggggcctgc gctacatggg aaagcttgtg 120
ggtgaggggc c 131

<210> 22
<211> 188
<212> DNA
<213> Homo sapiens

<400> 22
ctccttttag catctgctcc ctgcagggcc gtgccgctgt ccccacgtcg gctcacctgg 60
ccacctcacc tgcaggtagg aatcctaata ccaacctgga ggccctggaa gaatttaaga 120
aattggtgca gcacaaggga ctctcggagg aggacatttt cacgcccctg cagacgggtg 180
aggatcgg 188

<210> 23
<211> 44
<212> DNA
<213> Homo sapiens

<400> 23
gtctctgcag gaagctgcgt tctcgaacac tagggtgagt gagc 44

<210> 24
<211> 130
<212> DNA
<213> Homo sapiens

<400> 24
cccagcacag cagcccccgg gtctgcacct ccagagccca ccctaccacc agacacagag 60
cccggaccac ctggacctac cctccagcca tgacccttcc ctgctcccac ccacctgact 120
ccaaataaag 130

<210> 25
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<400> 25
Gly Thr Trp Tyr
1

<210> 26
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 26
cctgcaggta catgagcttc c 21

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 27
cgctgtattt gccaggctcc 20

<210> 28
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 28
gactcgagtc gacatcgatt ttttttttt ttttt 35

<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 29
cctctcccag ccccagcaag 20

<210> 30
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 30
gactcgagtc gacatcg 17

<210> 31
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 31
cgcccagtga cctgccgagg tc 22

<210> 32
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 32
ctttatttgg agtcaggtgg gtg 23

<210> 33
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 33
ctctgccccc tctgctgatg 20

<210> 34
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 34
cctgcttcac caccttcttg 20

<210> 35
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 35
gactcagact ccggagatga 20

<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 36
aactcagaca ccagagatga 20

<210> 37
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 37
gactcagatc ccggagatga 20

<210> 38
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 38
ccaggaggga ccactaca 18

<210> 39
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 39
ccgggacgga cgactacg 18

<210> 40
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 40
ctcatgacca cagtccatgc 20

<210> 41
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 41
tgttcgggaa cgcagctt 18

<210> 42
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 42
tgccgctgtc cccacgtcgg 20

<210> 43
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic motif

<400> 43
Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr Arg Leu Lys
1 5 10

<210> 44
<211> 28

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 44
gtcggatccc tgtccttcac cctggagg 28

<210> 45
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 45
gtcctcgagg tgttcgggaa cgcagcttc 29

<210> 46
<211> 176
<212> PRT
<213> Homo sapiens

<400> 46
Met Lys Pro Leu Leu Leu Ala Val Ser Leu Gly Leu Ile Ala Ala Leu
1 5 10 15

Gln Ala His His Leu Leu Ala Ser Asp Glu Glu Ile Gln Asp Val Ser
20 25 30

Gly Thr Trp Tyr Leu Lys Ala Met Thr Val Asp Arg Glu Phe Pro Glu
35 40 45

Met Asn Leu Glu Ser Val Thr Pro Met Thr Leu Thr Thr Leu Glu Gly
50 55 60

Gly Asn Leu Glu Ala Lys Val Thr Met Leu Ile Ser Gly Arg Cys Gln
65 70 75 80

Glu Val Lys Ala Val Leu Glu Lys Thr Asp Glu Pro Gly Lys Tyr Thr
85 90 95

Ala Asp Gly Gly Lys His Val Ala Tyr Ile Ile Arg Ser His Val Lys
100 105 110

Asp His Tyr Ile Phe Tyr Cys Glu Gly Glu Leu His Gly Lys Pro Val
115 120 125

Arg Gly Val Lys Leu Val Gly Arg Asp Pro Lys Asn Asn Leu Glu Ala
130 135 140

Leu Glu Asp Phe Glu Lys Ala Ala Gly Ala Arg Gly Leu Ser Thr Glu
145 150 155 160

```
Ser Ile Leu Ile Pro Arg Gln Ser Glu Thr Cys Ser Pro Gly Ser Asp
                165                 170                 175


<210> 47
<211> 176
<212> PRT
<213> Rattus sp.

<400> 47
Met Lys Ser Arg Leu Leu Thr Val Leu Leu Leu Gly Leu Met Ala Val
  1               5                  10                  15

Leu Lys Ala Gln Glu Ala Pro Pro Asp Asp Gln Glu Asp Phe Ser Gly
             20                  25                  30

Lys Trp Tyr Thr Lys Ala Thr Val Cys Asp Arg Asn His Thr Asp Gly
         35                  40                  45

Lys Arg Pro Met Lys Val Phe Pro Met Thr Val Thr Ala Leu Glu Gly
     50                  55                  60

Gly Asp Leu Glu Val Arg Ile Thr Phe Arg Gly Lys Gly His Cys His
 65                  70                  75                  80

Leu Arg Arg Ile Thr Met His Lys Thr Asp Glu Pro Gly Lys Tyr Thr
                85                  90                  95

Thr Phe Lys Gly Lys Lys Thr Phe Tyr Thr Lys Glu Ile Pro Val Lys
            100                 105                 110

Asp His Tyr Ile Phe Tyr Ile Lys Gly Gln Arg His Gly Lys Ser Tyr
        115                 120                 125

Leu Lys Gly Lys Leu Val Gly Arg Asp Ser Lys Asp Asn Pro Glu Ala
    130                 135                 140

Met Glu Glu Phe Lys Lys Phe Val Lys Ser Lys Gly Phe Arg Glu Glu
145                 150                 155                 160

Asn Ile Thr Val Pro Glu Leu Leu Asp Glu Cys Val Pro Gly Ser Asp
                165                 170                 175


<210> 48
<211> 178
<212> PRT
<213> Bovine sp.

<400> 48
Met Lys Cys Leu Leu Leu Ala Leu Ala Leu Thr Cys Gly Ala Gln Ala
  1               5                  10                  15

Leu Ile Val Thr Gln Thr Met Lys Gly Leu Asp Ile Gln Lys Val Ala
             20                  25                  30
```

```
Gly Thr Trp Tyr Ser Leu Ala Met Ala Ala Ser Asp Ile Ser Leu Leu
          35                  40                  45

Asp Ala Gln Ser Ala Pro Leu Arg Val Tyr Val Glu Glu Leu Lys Pro
      50                  55                  60

Thr Pro Glu Gly Asp Leu Glu Ile Leu Leu Gln Lys Trp Glu Asn Gly
 65                  70                  75                  80

Glu Cys Ala Gln Lys Lys Ile Ile Ala Glu Lys Thr Lys Ile Pro Ala
                 85                  90                  95

Val Phe Lys Ile Asp Ala Leu Asn Glu Asn Lys Val Leu Val Leu Asp
            100                 105                 110

Thr Asp Tyr Lys Lys Tyr Leu Leu Phe Cys Met Glu Asn Ser Ala Glu
        115                 120                 125

Pro Glu Gln Ser Leu Ala Cys Gln Cys Leu Val Arg Thr Pro Glu Val
    130                 135                 140

Asp Asp Glu Ala Leu Glu Lys Phe Asp Lys Ala Leu Lys Ala Leu Pro
145                 150                 155                 160

Met His Ile Arg Leu Ser Phe Asn Pro Thr Gln Leu Glu Glu Gln Cys
                165                 170                 175

His Ile
```

<210> 49
<211> 180
<212> PRT
<213> Murine sp.

<400> 49

```
Met Lys Met Leu Leu Leu Leu Cys Leu Gly Leu Thr Leu Val Cys Val
  1               5                  10                  15

His Ala Glu Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys
            20                  25                  30

Ile Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu
          35                  40                  45

Lys Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His
      50                  55                  60

Val Leu Glu Asn Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu
 65                  70                  75                  80

Glu Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly
                 85                  90                  95

Glu Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys
            100                 105                 110
```

```
Thr Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp
        115                 120                 125

Gly Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro Asp Leu
    130                 135                 140

Met Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Glu His Gly
145                 150                 155                 160

Ile Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg Cys Leu
                165                 170                 175

Gln Ala Arg Glu
            180
```

<210> 50
<211> 199
<212> PRT
<213> Homo sapiens

<400> 50

```
Met Lys Trp Val Trp Ala Leu Leu Leu Leu Ala Ala Trp Ala Ala Ala
  1               5                  10                  15

Glu Arg Asp Cys Arg Val Ser Ser Phe Arg Val Lys Glu Asn Phe Asp
            20                  25                  30

Lys Ala Arg Phe Ser Gly Thr Trp Tyr Ala Met Ala Lys Lys Asp Pro
        35                  40                  45

Glu Gly Leu Phe Leu Gln Asp Asn Ile Val Ala Glu Phe Ser Val Asp
    50                  55                  60

Glu Thr Gly Gln Met Ser Ala Thr Ala Lys Gly Arg Val Arg Leu Leu
65                  70                  75                  80

Asn Asn Trp Asp Val Cys Ala Asp Met Val Gly Thr Phe Thr Asp Thr
                85                  90                  95

Glu Asp Pro Ala Lys Phe Lys Met Lys Tyr Trp Gly Val Ala Ser Phe
            100                 105                 110

Leu Gln Lys Gly Asn Asp Asp His Trp Ile Val Asp Thr Asp Tyr Asp
        115                 120                 125

Thr Tyr Ala Val Gln Tyr Ser Cys Arg Leu Leu Asn Leu Asp Gly Thr
    130                 135                 140

Cys Ala Asp Ser Tyr Ser Phe Val Phe Ser Arg Asp Pro Asn Gly Leu
145                 150                 155                 160

Pro Pro Glu Ala Gln Lys Ile Val Arg Gln Arg Gln Glu Glu Leu Cys
                165                 170                 175

Leu Ala Arg Gln Tyr Arg Leu Ile Val His Asn Gly Tyr Cys Asp Gly
            180                 185                 190
```

```
Arg Ser Glu Arg Asn Leu Leu
        195
```

<210> 51
<211> 159
<212> PRT
<213> Bovine sp.

<400> 51

```
Ala Gln Glu Glu Glu Ala Glu Gln Asn Leu Ser Glu Leu Ser Gly Pro
  1               5                  10                  15

Trp Arg Thr Val Tyr Ile Gly Ser Thr Asn Pro Glu Lys Ile Gln Glu
             20                  25                  30

Asn Gly Pro Phe Arg Thr Tyr Phe Arg Glu Leu Val Phe Asp Asp Glu
         35                  40                  45

Lys Gly Thr Val Asp Phe Tyr Phe Ser Val Lys Arg Asp Gly Lys Trp
     50                  55                  60

Lys Asn Val His Val Lys Ala Thr Lys Gln Asp Asp Gly Thr Tyr Val
 65                  70                  75                  80

Ala Asp Tyr Glu Gly Gln Asn Val Phe Lys Ile Val Ser Leu Ser Arg
                85                  90                  95

Thr His Leu Val Ala His Asn Ile Asn Val Asp Lys His Gly Gln Thr
            100                 105                 110

Thr Glu Leu Thr Glu Leu Phe Val Lys Leu Asn Val Glu Asp Glu Asp
        115                 120                 125

Leu Glu Lys Phe Trp Lys Leu Thr Glu Asp Lys Gly Ile Asp Lys Lys
    130                 135                 140

Asn Val Val Asn Phe Leu Glu Asn Glu Asp His Pro His Pro Glu
145                 150                 155
```

<210> 52
<211> 181
<212> PRT
<213> Rattus sp.

<400> 52

```
Met Lys Leu Leu Leu Leu Leu Leu Cys Leu Gly Leu Thr Leu Val Cys
  1               5                  10                  15

Gly His Ala Glu Glu Ala Ser Ser Thr Arg Gly Asn Leu Asp Val Ala
             20                  25                  30

Lys Leu Asn Gly Asp Trp Phe Ser Ile Val Val Ala Ser Asn Lys Arg
         35                  40                  45

Glu Lys Ile Glu Glu Asn Gly Ser Met Arg Val Phe Met Gln His Ile
     50                  55                  60
```

```
Asp Val Leu Glu Asn Ser Leu Gly Phe Lys Phe Arg Ile Lys Glu Asn
 65                  70                  75                  80

Gly Glu Cys Arg Glu Leu Tyr Leu Val Ala Tyr Lys Thr Pro Glu Asp
                85                  90                  95

Gly Glu Tyr Phe Val Glu Tyr Asp Gly Gly Asn Thr Phe Thr Ile Leu
            100                 105                 110

Lys Thr Asp Tyr Asp Arg Tyr Val Met Phe His Leu Ile Asn Phe Lys
        115                 120                 125

Asn Gly Glu Thr Phe Gln Leu Met Val Leu Tyr Gly Arg Thr Lys Asp
    130                 135                 140

Leu Ser Ser Asp Ile Lys Glu Lys Phe Ala Lys Leu Cys Glu Ala His
145                 150                 155                 160

Gly Ile Thr Arg Asp Asn Ile Ile Asp Leu Thr Lys Thr Asp Arg Cys
                165                 170                 175

Leu Gln Ala Arg Gly
            180


<210> 53
<211> 157
<212> PRT
<213> Sus scrofa

<400> 53
Gln Glu Pro Gln Pro Glu Gln Asp Pro Phe Glu Leu Ser Gly Lys Trp
  1               5                  10                  15

Ile Thr Ser Tyr Ile Gly Ser Ser Asp Leu Glu Lys Ile Gly Glu Asn
             20                  25                  30

Ala Pro Phe Gln Val Phe Met Arg Ser Ile Glu Phe Asp Asp Lys Glu
         35                  40                  45

Ser Lys Val Tyr Leu Asn Phe Phe Ser Lys Glu Asn Gly Ile Cys Glu
     50                  55                  60

Glu Phe Ser Leu Ile Gly Thr Lys Gln Glu Gly Asn Thr Tyr Asp Val
 65                  70                  75                  80

Asn Tyr Ala Gly Asn Asn Lys Phe Val Val Ser Tyr Ala Ser Glu Thr
                85                  90                  95

Ala Leu Ile Ile Ser Asn Ile Asn Val Asp Glu Glu Gly Asp Lys Thr
            100                 105                 110

Ile Met Thr Gly Leu Leu Gly Lys Gly Thr Asp Ile Glu Asp Gln Asp
        115                 120                 125
```

```
Leu Glu Lys Phe Lys Glu Val Thr Arg Glu Asn Gly Ile Pro Glu Glu
    130                 135                 140

Asn Ile Val Asn Ile Ile Glu Arg Asp Asp Cys Pro Ala
145                 150                 155
```

HUMAN ODORANT-BINDING PROTEINS WHICH BIND HYDROPHOBIC LIGANDS: POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAID POLYPEPTIDES, AND APPLICATIONS THEREOF

The present invention relates to the demonstration of novel human odorant-binding proteins, hereafter named "OBPs", and also to the applications thereof, both therapeutically and non-therapeutically.

The present invention is based on the identification of a family of lipocalin genes, composed of three genes and two pseudogenes, on human chromosome 9q34; the three genes correspond to the LCN1 gene which has already been described and to two novel genes which are the subject of the present invention and are called hOBPIIa and hOBPIIb. In addition, the invention is based on the attribution of novel functions to already known human lipocalins, through the demonstration of new areas of expression.

Although a certain number of OBPs have already been demonstrated (Pelosi *et al.*, 1996), such as the rat OBP for example (see in particular patent EP-0 335 654), the OBPs according to the present invention are human OBPs which have a very large number of advantages, as will emerge from the remainder of the text, compared to the murine proteins.

These OBP proteins of the lipocalin family have, as far as some of them are concerned, been mentioned indirectly in patent WO 99/07740, but their OBP function has never been described to date; the same is also true of the proteins LCN1, retinol-binding protein (RBP) and apolipoprotein D (ApoD), as will be more fully explained hereinafter.

Historically, the lipocalin family (Pervaiz and Brew, 1987), has been defined based on the human retinol-

binding protein (RBP) and on 3 other proteins: bovine β-lactoglobulin, rat α2μ-globulin and human α1-microglobulin. From these reference proteins, and using sequence homologies, the lipocalin family has increased in size so as to now incorporate a large number of proteins, more than about 100, both in eukaryotes and in prokaryotes (Flower *et al.*, 1995, 1996). This family consists of small proteins (160-190 amino acids) which contain a hydrophobic pocket and are generally secreted (Bocskei *et al.*, 1992; Senoo *et al.*, 1990, Zeng *et al.*, 1996; Miller, 1998), although, in certain cases, they are proteins which remain associated with the membrane (Nagata *et al.*, 1991). In vertebrates, the sequence identities between the various lipocalins are in the region of 20%; however, the sequence identities are greater for the orthologous proteins, and also for the recent paralogous genes described (Igarashi *et al.*, 1992; Dewald *et al.*, 1992).

The present invention relates to the isolated OBPII polypeptides encoded by the two new isolated human genes, OBPIIa and OBPIIb, located at the 9q34 locus; the invention also relates to the corresponding polynucleotide sequences, the corresponding mRNAs and also the regulatory promoter sequences which determine the expression profile in the various tissues and in particular in the secretory tissues. These two genes encode at least seven different polypeptides given the existence of alternative splicing of the transcripts. The OBPIIa gene encodes at least 4 different polypeptides, named $OBPII_{a\alpha}$, $OBPII_{a\beta}$, $OBPII_{a\gamma}$ and $OBPII_{a\delta}$, and the OBPIIb gene encodes three different polypeptides, named $OBPII_{b\alpha}$, $OBPII_{b\beta}$ and $OBPII_{b\delta}$. Even though one gene is mainly expressed, all the messenger forms are found in the nasal structure.

The subject of the present invention is therefore an isolated polypeptide comprising an amino acid sequence having at least 90% identity with the amino acid

sequences SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 12 or SEQ ID No. 14, and LCN1, ApoD and RBP.

It should be understood that the invention relates to the polypeptides obtained by purification from natural sources or obtained using recombinant techniques, as will be described subsequently; the invention also relates to the polypeptides obtained by chemical synthesis, which may then comprise unnatural amino acids. In the present description, the term "polypeptide" will be used to denote equally a protein or a peptide.

The expression "percentage identity between the sequences" is intended to denote the percentage of amino acids which are identical between the sequences, obtained with the best possible sequence alignment, this percentage being purely statistical and the differences between the polypeptides being distributed randomly and over the entire length.

The present invention also relates to an isolated polypeptide, characterized in that it comprises a polypeptide chosen from:

a) a polypeptide of sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 12 or SEQ ID No. 14;

b) a polypeptide which is a variant of a polypeptide of amino acid sequences defined in a);

c) a polypeptide homologous to the polypeptide defined in a) or b) and comprising at least 90% identity with said polypeptide of a);

d) a fragment of at least 15 consecutive amino acids of a polypeptide defined in a), b) or c);

e) a biologically active fragment of a polypeptide defined in a), b) or c).

The term "variant polypeptide" will be intended to mean all the mutated polypeptides which may exist naturally, in particular in humans, and which correspond in particular to truncations, substitutions, deletions and/or additions of amino acid residues. The variant polypeptides according to the invention conserve at least one hydrophobic ligand-binding domain.

The term "homologous polypeptide" will be intended to denote the polypeptides exhibiting, compared to the polypeptides of SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 12 and SEQ ID No. 14, certain modifications, such as in particular a deletion, addition or substitution of at least one amino acid, a truncation, an extension and/or a chimeric fusion. Among the homologous polypeptides, preference is given to those in which the amino acid sequence exhibits at least 90% identity, in preference 95%, preferably 97%, and even more preferably 99% identity with the amino acid sequences of the polypeptides according to the invention. In the case of a substitution, one or more amino acids, which may or may not be consecutive, are replaced with "equivalent amino acids". The expression "equivalent amino acid" is herein aimed at denoting any amino acid which can be substituted for one of the amino acids of the basic structure, without however modifying the essential functional properties or characteristics of the corresponding polypeptides, their biological activities, such as for example the *in vivo* induction of antibodies capable of recognizing the polypeptide, the amino acid sequence of which is included in the amino acid sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 12 or SEQ ID No. 14, or a fragment thereof. These equivalent amino

acids may be determined either based on their structural homology with the amino acids for which they substitute, or based on the results of cross biological assays which the various polypeptides may produce. By way of example, mention will be made of the possibilities of substitutions which can be made without their being, as a result of this, an extensive modification of the biological activities of the correspondingly modified polypeptides; for example, replacing leucine with valine or isoleucine, aspartic acid with glutamic acid, glutamine with asparagine, arginine with lysine, etc., it naturally being possible to envision the reverse substitutions under the same conditions. Thus, it is possible to envision introducing certain modifications, such as in particular a deletion, addition or substitution of at least one amino acid, into the alpha-helices of the protein without destroying the calyx formed by the structure composed of the beta-sheets; similarly, it is possible to introduce equivalent amino acids which make it possible to conserve the hydrophobic nature of the beta-sheets. It may also be advantageous to introduce modifications into the sequence of the polypeptides of the invention in order to generate homologous polypeptides lacking protease sites.

The term "biologically active fragment" will be intended to denote, in particular, a fragment of amino acid sequence of a polypeptide according to the invention, exhibiting at least one of the functional properties or characteristics of the polypeptides according to the invention, in particular in that: (i) it is capable of being recognized by an antibody specific for a polypeptide according to the invention or by antibodies produced by patients during an immune reaction; (ii) it has at least one of the domains or regions as defined below; (iii) it is capable of binding a hydrophobic ligand and in particular odorous molecules, preferably pheromones; (iv) it is capable of specifically binding a receptor.

The term "polypeptide fragment" is intended to denote a polypeptide comprising a minimum of 15 amino acids, in preference 18 amino acids, preferably 25 amino acids and even more preferably 50 amino acids. The polypeptide fragments according to the invention, obtained by cleavage of said polypeptide with a proteolytic enzyme, or with a chemical reagent, or by placing said polypeptide in a very acid environment, are also part of the invention. When the intention is to use a sequence of 15, 18, 25 or 50 amino acids, this, of course, preferably involves the portions corresponding to functional epitopes of the above polypeptides, which may be integrated into a polypeptide with a longer structure, for example in the form of a fusion protein; this will depend on the intended uses of these polypeptides.

According to a preferred embodiment, the present invention relates to an isolated polypeptide selected from a polypeptide corresponding to the sequence SEQ ID No. 2 and named $OBPII_{a\alpha}$, to the sequence SEQ ID No. 4 and named $OBPII_{a\beta}$, to the sequence SEQ ID No. 6 and named $OBPII_{a\gamma}$, to the sequence SEQ ID No. 10 and named $OBPII_{b\alpha}$, and to the sequence SEQ ID No. 12 and named $OBPII_{b\beta}$.

The polypeptides according to the invention are characterized in that they preferably comprise the Gly-Thr-Trp-Tyr domain.

The invention also relates to the isolated polynucleotide characterized in that it encodes a polypeptide as described above. The polynucleotide according to the invention is chosen from the polynucleotide of sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 9, SEQ ID No. 11 or SEQ ID No. 13.

The invention also relates to an isolated polynucleotide, characterized in that it comprises a polynucleotide chosen from:

a) a polynucleotide of sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 9, SEQ ID No. 11 or SEQ ID No. 13, or the sequence of which is that of the RNA corresponding to the sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 9, SEQ ID No. 11 or SEQ ID No. 13;

b) a polynucleotide, the sequence of which is complementary to the sequence of a polynucleotide defined in a),

c) a polynucleotide, the sequence of which comprises at least 80% identity, in preference 90%, preferably 95%, and even more preferably 97% identity with a polynucleotide defined in a) or b),

d) a polynucleotide which hybridizes, under conditions of high stringency, with a sequence of a polynucleotide defined in a), b) or c),

e) a fragment of at least 15 consecutive nucleotides, in preference 21 consecutive nucleotides, and preferably 30 consecutive nucleotides, of a polynucleotide defined in a), b), c) or d).

In the present description, the terms "polynucleotide", "oligonucleotide", "polynucleotide sequence", "nucleotide sequence" or "nucleic acid" will be intended to denote a DNA fragment, both a double-stranded DNA, a single-stranded DNA and products for transcription of said DNAs, and/or an RNA fragment, said synthetic or isolated natural fragments, possibly comprising unnatural nucleotides, denoting a precise chain of nucleotides, which may or may not be modified, making

it possible to define a fragment or a region of a nucleic acid.

The expression "polynucleotide of complementary sequence" is intended to denote any DNA, the nucleotides of which are complementary to those of SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 9, SEQ ID No. 11 or SEQ ID No. 13, or a portion of SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 9, SEQ ID No. 11 or SEQ ID No. 13, and the orientation of which is reversed.

For the purpose of the present invention, the term "percentage identity" is intended to mean a percentage of bases which are identical between the polynucleotides, obtained after the best alignment, this percentage being purely statistical and the differences between the two polynucleotides being distributed randomly and over their entire length.

For the purpose of the present invention, the expression "hybridization under conditions of high stringency" means that the conditions of temperature and ionic strength are chosen such that they allow the hybridization between two complementary DNA fragments to be maintained. By way of illustration, conditions of high stringency for the hybridization step for the purposes of defining the polynucleotide fragments described above are advantageously as follows:

The DNA-DNA or DNA-RNA hybridization is carried out in two steps: (1) prehybridization at 42°C for 3 hours in phosphate buffer (20 mM, pH 7.5) containing 5 × SSC (1 × SSC corresponds to a solution of 0.15 M NaCl + 0.015 M sodium citrate), 50% of formamide, 7% of sodium dodecyl sulfate (SDS), 10 × Denhardt's, 5% of dextran sulfate and 100 µg/ml of salmon sperm DNA; (2) actual hybridization for 20 hours at a temperature which depends on the length of the probe (i.e.: 42°C for a

probe of length > 100 nucleotides), followed by 2 washes for 20 minutes at 20°C in 2 × SSC + 2% SDS, 1 wash for 20 minutes at 20°C in 0.1 × SSC + 0.1% SDS. The final wash is carried out in 0.1 × SSC + 0.1% SDS for 30 minutes at 60°C for a probe of length > 100 nucleotides. The high stringency hybridization conditions described above for a nucleotide of defined length will be adjusted by those skilled in the art for longer or shorter oligonucleotides, according to the teaching of Sambrook *et al.* (1989).

Advantageously, a nucleotide fragment corresponding to the above definition will have at least 15 consecutive nucleotides, preferably at least 21 nucleotides, and even more preferentially at least 30 consecutive nucleotides, of the sequence from which it is derived.

According to one embodiment of the invention, the polynucleotide according to the invention is characterized in that it is directly or indirectly labeled with a radioactive compound or a nonradioactive compound. The polynucleotide according to the invention is used as a primer for amplifying or polymerizing nucleic acid sequences; the invention also relates to the use of a polynucleotide according to the invention as a probe for detecting nucleic acid sequences. According to the invention, the polynucleotide fragments which may be used as a probe or as a primer in processes for detecting, identifying, titering or amplifying nucleic acid sequence will have a minimum length of 9 bases, preferably 18 bases, and even more preferably 36 bases. Finally, the invention relates to the use of a polynucleotide according to the invention as a sense or antisense oligonucleotide for controlling the expression of the corresponding protein product, in this case a polypeptide according to the invention.

The unlabeled polynucleotide sequences according to the invention may be used directly as a probe, primer or

oligonucleotide; however, the sequences used are generally labeled so as to obtain sequences which can be used for many applications. The primers, probes and oligonucleotides according to the invention are labeled with radioactive elements or with nonradioactive molecules; among the radioactive isotopes used, mention may be made of ^{32}P, ^{33}P, ^{35}S, ^{3}H or ^{125}I. The nonradioactive entities are selected from ligands, such as biotin, avidin, streptavidin or digoxigenin, haptens, dyes and luminescent agents, such as radioluminescent, chemiluminescent, bioluminescent, fluorescent or phosphorescent agents.

The invention also comprises a method for detecting and/or titering nucleic acid according to the invention, in a biological sample, characterized in that it comprises the following steps: (i) isolating the DNA from the biological sample to be analyzed, or obtaining a cDNA from the RNA of the biological sample; (ii) specifically amplifying the DNA encoding the polypeptide according to the invention using primers; (iii) analyzing the amplification products.

The invention also comprises a kit for detecting and/or titering a nucleic acid according to the invention, in a biological sample, characterized in that it comprises the following elements: (i) a pair of nucleic acid primers according to the invention, (ii) the reagents required to perform a DNA amplification reaction and, optionally, (iii) a component for verifying the sequence of the amplified fragment, more particularly a probe according to the invention.

The invention also comprises a method for detecting and/or titering nucleic acid according to the invention, in a biological sample, characterized in that it comprises the following steps: (i) bringing a probe according to the invention into contact with a biological sample; (ii) detecting and/or titering the

hybrid formed between said probe and the DNA of the biological sample.

The invention also comprises a kit for detecting and/or titering nucleic acid according to the invention, in a biological sample, characterized in that it comprises the following elements: (i) a probe according to the invention, (ii) the reagents required to carry out a hybridization reaction and, where appropriate, (iii) a pair of primers according to the invention, and also the reagents required for a DNA amplification reaction.

The invention in particular relates to the methods according to the invention described above, for detecting and diagnosing cells of cancerous origin, and mainly breast, uterine, ovarian, prostate and lung cancers.

The polynucleotides according to the invention may thus be used as a primer and/or probe in processes using in particular the PCR (polymerase chain reaction) technique (Erlich, 1989; Innis *et al.*, 1990, and Rolfs *et al.*, 1991). This technique requires choosing pairs of oligonucleotide primers which border the fragment which must be amplified. Reference may, for example, be made to the technique described in American patent U.S. No. 4 683 202. The amplified fragments can be identified, for example after agarose or polyacrylamide gel electrophoresis or after a chromatography technique such as gel filtration or ion-exchange chromatography. The specificity of the amplification can be controlled by molecular hybridization using, as a probe, the nucleotide sequences polynucleotides of the invention, plasmids containing these sequences or products of amplification thereof. The amplified nucleotide fragments may be used as reagents in hybridization reactions in order to demonstrate the presence, in a biological sample, of a target nucleic acid of sequence

complementary to that of said amplified nucleotide fragments.

The invention is also directed toward the nucleotide fragments which can be obtained by amplification using primers according to the invention.

Other techniques for amplifying target nucleic acid may advantageously be used as an alternative to PCR (PCR-like), using pairs of primers of nucleotide sequences according to the invention. The term "PCR-like" will be intended to denote all methods using direct or indirect reproductions of nucleic acid sequences, or in which the labeling systems have been amplified; these techniques are of course known; in general, this involves amplifying the DNA with a polymerase; when the sample of origin is an RNA, a reverse transcription should be performed beforehand. There are currently a large number of processes for this amplification, such as for example the SDA (strand displacement amplification) technique (Walker *et al.*, 1992), the TAS (transcription-based amplification system) technique described by Kwoh *et al.* in 1989, the 3SR (self-sustained sequence replication) technique described by Guatelli *et al.* in 1990, the NASBA (nucleic acid sequence based amplification) technique described by Kievitis *et al.* in 1991, the TMA (transcription mediated amplification) technique, the LCR (ligase chain reaction) technique described by Landegren *et al.* in 1988 and improved by Barany *et al.* in 1991, which uses a heat-stable ligase, the RCR (repair chain reaction) technique described by Segev in 1992, the CPR (cycling probe reaction) technique described by Duck *et al.* in 1990, and the Q-beta-replicase amplification technique described by Miele *et al.* in 1983 and improved in particular by Chu *et al.* in 1986 and Lizardi *et al.* in 1988, and then by Burg *et al.* (1996) and also by Stone *et al.* (1996).

When the target polynucleotide to be detected is an RNA, for example an mRNA, use will advantageously be made, prior to carrying out an amplification reaction using the primers according to the invention or to carrying out a detection process using the probes of the invention, of an enzyme of the reverse transcriptase type in order to obtain a cDNA from the RNA contained in the biological sample. The cDNA obtained will then be used as a target for the primers or the probes used in the amplification or detection process according to the invention.

The nucleotide probes according to the invention hybridize specifically with a DNA or RNA molecule of a polynucleotide according to the invention, more particularly with the sequences SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 9, SEQ ID No. 11 and SEQ ID No. 13, under high stringency hybridization conditions as given in the form of a previous example.

The hybridization technique may be carried out in diverse ways (Matthews *et al.*, 1988). The most general method consists in immobilizing the nucleic acid extracted from the cells of various tissues or from cells in culture, on a support (such as nitrocellulose, nylon or polystyrene), and in incubating the immobilized target nucleic acid with the probe, under well-defined conditions. After hybridization, the excess probe is removed and the hybrid molecules formed are detected by the suitable method (measurement of the radioactivity, of the fluorescence or of the enzymatic activity linked to the probe).

According to another embodiment of the nucleic acid probes according to the invention, they may be used as a capture probe. In this case, a probe, termed "capture probe", is immobilized on a support and is used to capture, by specific hybridization, the target nucleic

acid obtained from the biological sample to be tested, and the target nucleic acid is then detected using a second probe, termed "detection probe", labeled with an easily detectable element.

In a preferred embodiment, the invention comprises the use of a sense or antisense oligonucleotide for controlling the expression of the corresponding protein product. Among the advantageous nucleic acid fragments, mention should in particular be made of the antisense oligonucleotides the structure of which ensures, by hybridization with the target sequence, inhibition of the expression of the corresponding product. Mention should also be made of the sense oligonucleotides which, by interaction with proteins involved in regulating the expression of the corresponding product, will induce either inhibition or activation of this expression. The oligonucleotides according to the invention have a minimum length of 9 bases, preferably of 18 bases, and even more preferably 36 bases.

The invention relates to a recombinant vector for cloning a polynucleotide according to the invention and/or for expressing a polynucleotide according to the invention, characterized in that it contains a polynucleotide according to the invention as described above. The vector according to the invention is characterized in that it comprises the elements which allow the expression of said sequences in a host cell and, optionally, the secretion of said sequences out of the host cell.

The term "expression vector" is intended to mean both autonomously replicating expression vectors of the plasmid type and systems intended to ensure integration in the cells, but these expression vectors may also be expression vectors of the viral type or even, when the intention is to perform, for example, gene therapy, naked DNA. Among the viral vectors, preference is given

to those derived from adenoviruses, from adenovirus-associated viruses (AAVs), from retroviruses, from lentiviruses, and preferably HIV derivatives, from poxviruses and from the herpesvirus, for expression in a eukaryotic system. Among the nonviral vectors, preference is given to naked polynucleotides, such as naked DNA or naked RNA according to the technique developed by the company VICAL, yeast artificial chromosomes (YACs) for expression in yeast, mouse artificial chromosomes (MACs) for expression in murine cells and, preferably, human artificial chromosomes (HACs) for expression in human cells.

According to a particular embodiment, the vector according to the invention comprises elements for controlling the expression of the polypeptides; these control elements are preferably chosen from (i) the promoter sequence of the hOBPIIa gene according to the invention, which corresponds to the sequence SEQ ID No. 15 and/or from the promoter sequence of the hOBPIIb gene according to the invention, which corresponds to the sequence SEQ ID No. 16; (ii) a polynucleotide, the sequence of which is complementary to the sequence SEQ ID No. 15 and SEQ ID No. 16; (iii) a polynucleotide, the sequence of which comprises at least 80% identity with a polynucleotide defined in (i) or (ii); (iv) a polynucleotide which hybridizes, under conditions of high stringency, with the polynucleotide sequence defined in (i), (ii) or (iii).

The computer tools available to those skilled in the art allow them to easily identify the regulatory promoter boxes required and sufficient to control the gene expression, in particular the TATA, CCAAT and GC boxes, and also the stimulatory (enhancer) or inhibitory (silencer) regulatory sequences which control, in CIS, the expression of the genes according to the invention; among these regulatory sequences, mention should be made of IRE, MRE and CRE.

It is also within the scope of the invention to use the elements defined above and chosen from the sequence SEQ ID No. 15 and SEQ ID No. 16, for controlling the expression of heterologous polypeptides other than those of the invention, and in particular for directing the expression of heterologous polypeptides in the cell types in which the polypeptides according to the invention are normally expressed.

The invention also comprises the host cells, in particular the eukaryotic and prokaryotic cells, characterized in that they are transformed with the vectors according to the invention. The host cells are preferably transformed under conditions which allow the expression of a recombinant polypeptide according to the invention. The host cell may be chosen from bacterial cells, but also from yeast cells and from plant and animal cells; the host cell is preferably a mammalian cell (Edwards and Aruffo, 1993), but also an insect cell in which it is possible to use processes which use baculoviruses for example (Luckow, 1993). These cells may be obtained by introducing into host cells a nucleotide sequence inserted into a vector as defined above, and then culturing said cells under conditions which allow the replication and/or expression of the transfected nucleotide sequence.

According to a particular embodiment, the invention also relates to a transgenic animal or plant which comprises host cells according to the invention.

The invention also relates to a method for preparing a polypeptide, characterized in that it uses a vector according to the invention. More particularly, the invention relates to a method for preparing a recombinant polypeptide, characterized in that transformed cells according to the invention are cultured under conditions which allow the expression of said

recombinant polypeptide, and in that said recombinant polypeptide is recovered.

The polypeptide according to the invention can be obtained according to a process of the invention and according to the techniques for producing recombinant polypeptides known to those skilled in the art. The present invention therefore relates to the recombinant polypeptide which can be obtained using the method presented above. In this case, the nucleic acid sequence used is placed under the control of signals which allow its expression in a host cell. An effective system for producing a recombinant polypeptide requires a vector, for example of plasmid or viral origin, and a compatible host cell. The vector must comprise a promoter, translation initiation and termination signals, and also suitable regions for regulating transcription. It must be possible for it to be maintained stably in the cell and it may, optionally, contain particular signals which specify secretion of the translated polypeptide. These various control signals are chosen as a function of the host cell used. To this effect, the nucleic acid sequences according to the invention may be inserted into vectors which replicate autonomously in the host chosen, or vectors which integrate in the host chosen. Such vectors will be prepared according to the methods commonly used by those skilled in the art, and the clones resulting therefrom may be introduced into a suitable host using standard methods, such as for example transfection by calcium phosphate precipitation, lipofection, electroporation or heat shock.

The recombinant polypeptides obtained as indicated above may be both in the glycosylated form and nonglycosylated form, and may or may not have the natural tertiary structure.

The polypeptides obtained by chemical synthesis and possibly comprising unnatural amino acids, corresponding to said recombinant polypeptides, are also included in the invention. The peptides according to the invention may also be prepared using techniques which are conventional in the domain of peptide synthesis. This synthesis may be performed in homogeneous solution or in solid phase.

The recombinant polypeptide purification processes used are known to those skilled in the art. The recombinant polypeptide may be purified from cell lysates and extracts, and/or from the culture medium supernatant, by methods used individually or in combination, such as fractionation, chromatography methods, immunoaffinity techniques using specific monoclonal or polyclonal antibodies, etc. A preferred variant consists in producing a recombinant polypeptide fused to a "carrier" protein (chimeric protein). The advantage of this system is that it allows stabilization and a decrease in the proteolysis of the recombinant product, an increase in solubility during the *in vitro* renaturation and/or simplification of the purification when the fusion partner has affinity for a specific ligand.

The present invention relates to the use of a polypeptide chosen from an hOBPIIa and hOBPIIb polypeptide according to the invention, or a fragment thereof, and from LCN1, retinol-binding protein (RBP) and apolipoprotein D (ApoD), as a protein which binds to a hydrophobic ligand, preferably an odorous molecule. In fact, besides the OBPII polypeptides mentioned above, the OBP function of the lipocalins LCN1, RBP and ApoD has never been suggested. Specifically, while the presence of LCN1 in the nasal mucus has been described (Redl et al., 1992), it was, in fact, attributed another function, namely an antiprotease function, whereas the main function of the LCN1 protein purely

appears to be the transport of hydrophobic ligands; moreover, the tests carried out in the context of the present invention with a recombinant LCN1 protein have not made it possible to demonstrate antiprotease activity. The same is true for retinol-binding protein and apolipoprotein D, which have never been presented as having OBP function.

The present invention also relates to the use of a polypeptide according to the invention or a fragment thereof, LCN1, RBP and apolipoprotein D as a competitive inhibitor, as an agonist or as an antagonist of the cellular receptors for lipocalins. The use of an inhibitor of the binding of lipocalins to their receptor may be employed in an anticancer strategy; specifically, a certain number of tumors are hormone-dependent: thus, breast tumors are sensitive to steroids; since the steroids are transported by lipocalins, it is within the scope of the invention to provide inhibitors of binding to the polypeptides according to the invention and LCN1, OBP or APO-D to avoid the binding of steroid hormones to the receptors of tumor cells.

The invention also relates to a monoclonal or polyclonal antibody and fragments thereof, characterized in that they specifically bind a polypeptide according to the invention. Chimeric antibodies, humanized antibodies and single-chain antibodies are also part of the invention. The antibody fragments according to the invention are preferably Fab or F(ab')2 fragments.

The polypeptides according to the invention make it possible to prepare monoclonal or polyclonal antibodies. The monoclonal antibodies may advantageously be prepared from hybridomas according to the technique described by Kohler and Milstein in 1975. The polyclonal antibodies may be prepared for example by immunizing an animal, in particular a mouse, with a

polypeptide according to the invention combined with an adjuvant of the immune response, and then purifying the specific antibodies contained in the serum of the immunized animals on an affinity column to which the polypeptide having been used as the antigen has been attached beforehand. The polyclonal antibodies according to the invention may also be prepared by purification on an affinity column on which a polypeptide according to the invention has been immobilized beforehand.

The invention also relates to a monoclonal antibody specific for a polypeptide according to the invention and capable of inhibiting the interaction between said polypeptide and the cellular receptor to which said polypeptide specifically binds. According to another embodiment, the monoclonal antibody according to the invention is capable of inhibiting the interaction between said polypeptide and its hydrophobic ligands, in preference the odorous molecules, and preferably the pheromones, to which said polypeptide binds.

The antibodies of the invention may also be labeled in the same way as described previously for the nucleic acid probes of the invention, and preferably with labeling of the enzymatic, fluorescent or radioactive type. Such labeled antibodies may be used to detect these polypeptides in a biological sample. The biological sample preferably consists of a fluid, for example human biopsies, serum or blood. They thus constitute a means for analyzing the expression of the polypeptide according to the invention, for example by immunofluorescence and/or gold-labeling, and/or with enzymatic immunoconjugates.

More generally, the antibodies of the invention may advantageously be used in any situation in which the expression of a polypeptide according to the invention is to be observed, and more particularly in immuno-cytochemistry, in immunohistochemistry or in Western

blotting experiments, or in ELISA and RIA techniques. It is thus within the scope of the invention to provide a method for detecting and/or titering a polypeptide according to the invention, in a biological sample, characterized in that it comprises the steps of bringing the biological sample into contact with antibodies according to the invention, and then revealing the antigen-antibody complex formed.

Also falling within the context of the invention is a kit for detecting and/or titering a polypeptide according to the invention, in a biological sample, characterized in that it comprises the following elements: (i) a monoclonal or polyclonal antibody as described above; (ii) where appropriate, the reagents for constituting the medium suitable for the immunological reaction; (iii) the reagents for detecting the antigen-antibody complexes produced by the immunological reaction. This kit is in particular useful for carrying out Western blotting experiments and for immunoprecipitation experiments.

The hOBPIIs according to the present invention may be used in many applications.

The first applications of the OBPs lie in the control of odorants; these applications essentially concern personal hygiene (perfumery, cosmetology, pharmacy) or general hygiene.

They may be used in a process for controlling the volatilization of an odorant; such a process is characterized in that it comprises a step of binding of said odorant with a polypeptide according to the invention or with the LCN1, RBP or apolipoprotein D proteins.

It is also possible to bind the odorant to a solid support using an OBPII according to the invention

attached to said support, it being possible for this binding to be both covalent binding and noncovalent binding, for example by adsorption or by binding of the avidin-biotin type, if necessary. Under these conditions, a fragranced support is obtained which retains the odorants for a longer period of time, gradually releasing them and possibly being used both in the cosmetics domain and in the domain of maintenance products in general. The supports used may be supports of the block type or of the ball type, for example. Of course, the OBPIIs according to the present invention may be more particularly used in the perfumery and cosmetics industry, in which the OBPs will be used in the form of a liquid mixture intended to control the volatilization of the odorants after the composition has been spread on human skin. This makes it possible to prolong the "staying-power" of fragrances or to form part of the composition of body deodorants in particular. Of course, the term "fragrance" generally encompasses the mixtures known in perfumery, based on alcohol or in aqueous form, and containing in particular essential oils. The polypeptide according to the invention may also form part of the composition of retinol-based creams, as an agent for transporting and protecting said retinol, for cosmetological applications, and in particular for preventing, smoothing out and treating wrinkles and fine lines of the skin, and combating cutaneous and/or subcutaneous loosening.

Among the applications relating to general hygiene, the process according to the invention may be used in a device directed toward deodorizing premises, such as for example animal houses or cowsheds. Such a device may also be envisioned for deodorizing the flow of air entering an air conditioning machine; such a device would be very useful in geographical regions which are polluted.

The invention also relates to a process for screening a molecule, preferably odorants or flavors, which comprises passing the molecule over a substrate which comprises a polypeptide according to the invention or LCN1, RBP and apolipoprotein D, bound to said substrate, said polypeptide binding said odorant or flavor, and recovering said odorant or flavor from the polypeptide if necessary. The OBPs according to the invention may thus make it possible to isolate odorants or fragrances, by binding them to a support as described above but which, in this case, may be, for example, a chromatography column, by passing over this column the product from which it is desired to capture the odorants. It is, of course, possible, still in this same type of application, to carry out an analysis of a complex fragrance using the differences in retention of the various products by the OBPs.

The isolation process described above may also be used to separate particular "odorants", especially human pheromones.

The hOBPIIs according to the present invention may be used for agrofoods or industrial applications.

The OBPs according to the present invention, by virtue of their characteristics, may make it possible to solubilize certain lipophilic molecules by associating them with said OBPs. Thus, the polypeptides according to the invention may be used in combination with dietary fatty acids by way of a food supplement. The present invention relates to a process for solubilizing lipophilic molecules, characterized in that it comprises binding said lipophilic molecule to a polypeptide according to the invention, to LCN1, to RBP or to apoD.

The OBPIIs, like LCN1, RBP and ApoD, may be involved in the transport of fatty acids and in the biological

mechanisms for detecting the fatty acid load of the food intake, in particular in the buccal region. The invention therefore also relates to the use of the polypeptides above, in combination with fatty acids, for decreasing the consumption of fatty acids, in particular in hyperlipidemias or in obesity. The polypeptides above may therefore be used for treating hyperlipidemias and obesity. Specifically, these proteins contribute to the detection of the fatty acid content in the food taken in (Gilbertson, 1998); an excess of these proteins should lead to the physiological system which detects the fatty load of a food intake being deceived. Thus, a food portion which is low in fat but supplemented with OBP, ApoD, RBP or LCN1 pre-loaded with fatty acid will be falsely identified as being rich in fat.

Another application is to supplement a nonmaternal milk with one or more of the polypeptides described above.

The OBPs according to the present invention may be used in preventive and curative therapy.

Thus, a lack of detection of this type of protein in a biological specimen using an antibody, a primer and/or a probe according to the invention may be an element for the diagnosis of anosmia.

Similarly, it is possible to envision the detection of anti-OBPII antibodies in a biological sample of human origin; the presence or the titer of these antibodies may be related to certain types of allergy, in particular in asthmatics. In this case, it may be possible to treat this type of allergy by administering polypeptide fragments as described, which would decrease the immune reactions to the external allergies.

The invention therefore relates to a process for detecting antibodies against human OBPIIs (hOBPIIs) in human serum from an allergic and/or asthmatic patient using an hOBPII polypeptide according to the invention.

According to another embodiment, the invention relates to a process for detecting antibodies against human OBPIIs (hOBPIIs) in a biological fluid from a patient suffering from a cancer, and in particular from prostate cancer and/or from breast cancer and/or from uterine cancer and/or from ovarian cancer and/or from lung cancer. Specifically, the proteins according to the invention are overexpressed in tumors, and in particular in prostate cancer (US 5804368, WO 97 10503, CS 84 02898, CS 83 02012, CS 82 08506, CS 82 08215), breast cancer (Stoecz and Gould, 1995; Simard *et al.*, 1992), uterine cancer, ovarian cancer and lung cancer.

The OBPs according to the present invention may also be used in pharmaceutical compositions, this being in particular in order to deliver certain drugs.

In fact, lipocalins are used by mammals to transport hydrophobic molecules in biological fluids. It even appears that they are the natural transporters of xenobiotics *in vivo*. For example, an experimental overload of xenobiotics creates, in male rats, tumors in the proximal convoluted tube (PCT) of the nephron (Borghoff *et al.*, 1990). Specifically, the MUP proteins, only produced in male rats, are reabsorbed by the PCT cells; after lysosomal degradation, the MUP proteins release the xenobiotics which they were transporting. These xenobiotics accumulate in these cells and lead them toward a tumor pathway by mutagenesis. Lipocalins therefore appear to be the ideal structure to trap, in their calyx, a molecule used as a medicinal product, thus avoiding a general distribution thereof. Thus, G. Beste *et al.* (1999) and application WO 99 16873 describe a strategy of mutagenesis and screening

variants of a lipocalin having optimum affinity for a given ligand. It is therefore possible to create a highly specific "medicinal product cage".

The present invention therefore relates to a polypeptide according to the invention, as an agent for targeting a pharmaceutical compound. The expression "agent for targeting" is intended to denote the polypeptides of the invention capable of transporting and releasing the ligand to which they are bound, at the level of certain target tissues or cells which have, at their surface, receptors for said polypeptides. Thus, the invention relates to the transport of a medicinal product in the cage of a polypeptide according to the invention, and targeting cells through the ability of the OBPII to bind to a specific receptor.

It is also within the scope of the invention to develop fusion proteins or chimeric proteins comprising the cage corresponding to the lipocalins according to the invention combined with a protein which allows specific cellular addressing, generally different from that which is natural for the OBPIIs. A subject of the present invention is therefore a polypeptide, characterized in that said polypeptide is expressed in the form of a protein from fusion with a protein which allows specific cellular addressing. Among the proteins which allow specific cellular addressing, mention should be made of interleukins, of cytokines, of lymphokines, of chemokines, of growth factors, of hormones, and of monoclonal or polyclonal antibodies. The interleukins, cytokines and lymphokines are chosen from a group preferably composed of the interleukins II-1 to II-20, and the interferons α-IFN, β-IFN and γ-IFN. The growth factors are preferably chosen from the colony stimulating factors G-CSF, GM-CSF, M-CSF, and erythropoietin. The hormones are preferably chosen from steroid hormones.

According to another embodiment, the polypeptide according to the invention, as an agent for targeting a pharmaceutical compound, is combined with a molecule which allows specific cellular addressing. Among the molecules which allow specific cellular addressing, mention should be made of the group of steroids, of interleukins, of cytokines, of lymphokines, of interferons, of growth factors, of hormones and of antibodies. The molecule is preferably a steroid. The combination of the polypeptide according to the invention and said molecule is produced either via a noncovalent attachment, using for example the avidin-biotin system, or via a covalent attachment, using for example chemical bridging agents.

Among the pharmaceutical compounds which the polypeptide according to the invention may transport and target, mention should be made of medicinal products, and in particular anticancer agents. The anticancer agents are selected from the group of antiproliferative, antineoplastic or cytotoxic agents, and are used to arrest the development of cancers and to induce regression and/or elimination of the tumor mass. These anticancer agents are preferably radioisotopes, and even more preferably gamma-ray-emitting radioisotopes, such as iodine^{131}, yttrium^{90}, gold^{199}, palladium^{100}, copper^{67}, bismuth^{217} and antimony^{211}. Beta- and alpha-ray-emitting radioisotopes may also be used for the therapy. The non-isotopic anticancer agents bound to the polypeptide according to the invention are many and varied; mention may be made of: (i) antimetabolites, such as antifolate agents, methotrexate, (ii) purine and pyrimidine analogs (mercaptopurine, fluorouracil, 5-azacytidine), (iii) antibiotics, (iv) lectins (ricin, abrin) and (iv) bacterial toxins (diphtheria toxin); the toxins are preferably chosen from *Pseudomonas* exotoxin A, diphtheria toxin, cholera toxin, *Bacillus* anthrox toxin, Pertussis toxin, *Shigella* Shiga toxin,

Shiga-like toxin, *Escherichia coli* toxins, colicin A, d-endotoxin and *Haemophilus A* hemagglutinin.

The invention also relates to a pharmaceutical composition comprising a pharmaceutical compound bound to at least one polypeptide according to the invention, and a pharmaceutically acceptable vehicle. It is also within the scope of the invention to increase the binding capacity by mutagenesis of this same protein. The invention also relates to a pharmaceutical composition as defined above, for treating cancer, and in particular prostate cancer, breast cancer, uterine cancer, ovarian cancer, liver cancer and pulmonary epithelial cell carcinoma. Given that the OBP named hOBPIIb, apolipoprotein D, RBP and α-1-acid glycoprotein are expressed in the prostate, the pharmaceutical composition according to the invention is intended for the treatment of prostate cancer. Similarly, given that it has been demonstrated that the mRNA of the ApoD, RBP and LCN1 proteins is produced in the mammary gland, in addition to the production of hOBPIIb already described in application WO 99 07740, the pharmaceutical composition according to the invention is intended for the treatment of breast cancer.

The invention also relates to a polypeptide according to the invention, as a pharmaceutical compound transporter. The term "transporter" is intended to denote polypeptides according to the invention capable of vehicling a pharmaceutical compound, in the body, without said compound being released at a preferred site in the body. Such a polypeptide constitutes a means for delivering said pharmaceutical compound in the body.

The present invention also relates to a pharmaceutical composition according to the invention, characterized

in that said polypeptide constitutes a delayed form of delivery of said pharmaceutical compound in the body.

Specifically, the pharmaceutical compound of interest bound to the OBPII polypeptide according to the invention gradually diffuses in the body as said transporter polypeptide is catabolized in the body. The use of recombinant DNA techniques allows those skilled in the art to modify the half-life duration of the OBPII polypeptide in the body by introducing modifications into said polypeptide. Thus, it may be advantageous to develop a homologous polypeptide according to the invention, for which the protease cleavage sites have been mutated in order to increase the half-life of said polypeptide in the body. It may also be advantageous to develop a multimeric polypeptide, expressed for example in the form of a fusion protein, in order to avoid a "glomerular escape" (renal clearance) and thus to increase the half-life of said polypeptide in the body.

According to another aspect, the invention relates to a compound characterized in that it is chosen from an antibody, a polypeptide, a ligand, a polynucleotide, an oligonucleotide and a vector according to the invention, as a medicinal product, and in particular as active principles of a medicinal product; these compounds will preferentially be in soluble form, and combined with a pharmaceutically acceptable vehicle. The term "pharmaceutically acceptable vehicle" is intended to denote any type of vehicle conventionally used in the preparation of injectable compositions, i.e. a diluent and/or a suspending agent, such as an isotonic or buffered saline solution. These compounds will preferably be administered systemically, in particular intravenously, intramuscularly, intradermally or orally. Their optimal methods of administration, doses and pharmaceutical forms may be determined according to the criteria generally taken into account

in establishing a treatment suitable for a patient, such as for example the age or bodyweight of the patient, the seriousness of his or her general condition, the tolerance to the treatment and the side effects observed, etc.

The invention also relates to a pharmaceutical composition comprising an expression vector according to the invention and a pharmaceutically acceptable vehicle.

It is also possible to use the property of expression in particular sites of these proteins in order to provide imaging and/or a diagnosis of certain types of cancer.

Finally, given the high expression of the polypeptides according to the invention in the placenta (cf. figure 10), the present invention provides various therapeutic applications intended to treat pathological conditions of gestation in women.

Thus, one of the objects of the present invention is to use the polypeptide according to the invention for preparing a medicinal product intended to transport compounds across the placental barrier. These compounds are preferably lipophilic molecules. The polypeptide according to the invention may be used to transport compounds from the gestating mother to the fetus; the compounds possibly being chosen from the molecules of therapeutic interest, which are preferably lipophilic, essential fatty acids, hormones, vitamins and, more generally, small lipophilic molecules.

In fact, it has been demonstrated that the cells of the placental villi secrete massive amounts of the polypeptide according to the invention, in particular the OBPIIs. These polypeptides then spread through the

fetal blood and to the general blood circulation of the fetus.

In addition, the polypeptide according to the invention may be used to transport a compound from the fetus to the gestating mother. Specifically, the fetus secretes, via the villi of the nasal structures, polypeptides according to the invention which participate in molecule exchanges between the fetus and the mother at the amniotic membrane. An object of the present invention is therefore also to provide the fetus with polypeptides according to the invention so as to allow the transport of compounds present in the fetus, for example xenobiotics, to the placenta and/or to the amniotic membrane, in order to allow the detoxification of the fetus.

According to another aspect of the invention, the polypeptide according to the invention may be used as a pregnancy marker. Specifically, the polypeptide according to the invention is present in various biological liquids of the mother, such as the urine or blood; measuring the concentration thereof and/or the presence thereof constitutes a pregnancy marker. In this capacity, this marker may be used with, or in place of, or in addition to the pregnancy marker constituted by human choriogonadotrophin, hCG; the polypeptide according to the invention and hCG are, in fact, molecules which are secreted by the trophoblastic placental cells.

Finally, according to another aspect of the invention, the polypeptide may be used as a marker for feto-placental pathological conditions. It may constitute a marker for the rupture of the amniotic sack membrane.

Other characteristics and advantages of the invention appear in the remainder of the description with the

examples and the figures, the legends of which are represented below.

FIGURE 1: Genomic organization of the LCN1/hOBPII locus

The top line represents the region of chromosome 9q34. The double-headed arrow indicates the gap between the location of the polymorphic markers D9S1811 and D9S67; the relative position of the markers D9S67 and D9S1826 is uncertain.

The middle level indicates the partial organization of the cosmids at 3 different loci: LCN1c, LCN1b-$hOBP_{IIb}$, LCN1-$hOBP_{IIa}$. The cosmids (approximately 40 kb) are represented by horizontal lines with their name and the arms of the T3 and T7 vectors noted below the line. The arrows represent the various genes or pseudogenes with their respective orientation. In the cosmids, the vertical dotted lines represent the EcoRI sites. The symbol ⊗ indicates an uncertain orientation of the locus.

The bottom level shows the intron/exon structures of the LCN1 and hOBPII genes: the black boxes represent the exons; the arrows represent the transcription initiation site; o11 to o15 represent the oligonucleotide probes used to screen the cosmids; *Alu* represents the presence of repeat sequences.

FIGURE 2: Dot Plot Analysis of:

(A) sequence of the LCN1-hOBPIIa locus (AC000396+ACXXXXX) versus the LCN1b-hOBPIIb locus (AC002098);

(B) sequence of the LCN1-hOBPIIa locus (AC000396+ACXXXXX) versus the sequence of the LCN1c locus;

(C) sequence of the LCN1b-hOBPIIb locus (AC002098) versus the sequence of the LCN1c locus (AC002106).

Genomic sequences were filtered for the repetitive sequences using "Repeatmasker" (REF), and then compared

and dotplotted with the program gcg using a window size of 25, and a criterion of 20 with an 80% homology.

FIGURE 3: Nucleotide sequence of hOBPII genes

The upper lines represent the hOBPIIa sequence and the lower lines represent the hOBPIIb sequence for which only the different nucleotides are represented; a dash indicates the absence of corresponding sequences. The shaded capitals are the exon sequences and the lower case letters are the intron sequences. The sizes indicated on the left are indicated in bp. The TATA box is in bold characters and the polyadenylation signal is underlined. The boxes indicate the splice acceptor sites for exons 5, 5b and 5c.

FIGURE 4: Diagrammatic representation of the two hOBPII genes and of their corresponding mRNAs

The horizontal lines represent the exon/intron organization with sizes indicated in bp. The shaded boxes numbered from 1 to 7 are the coding exon sequences of the main transcripts; the letters b and c refer to supernumerary exons. The various transcripts are represented using assembled boxes: the first, hOBPIIaα and hOBPIIbα, correspond to the main transcripts, the others correspond to the forms derived from alternate splicing. ↓ indicates a shift in the reading frame resulting from the insertion or from the deletion of an exon and * represents a stop codon. The letter "a" represents an α-helix and "b" β-sheets predicted by the DSC program. The letters in italics are predictions obtained with the "Predator" program.

FIGURE 5: Alignment of the protein sequences derived from the two human genes hOBPIIa and hOBPIIb (hOBPIIaα = OBP2aaHOMS, hOBPIIbα = OBP2baHOMSA, hOBPIIbβ = OPB2bbHOMSA, hOBPIIaγ = OBP2agHOMSA, hOBPIIaβ = OBP2abHOMSA), of human tear lipocalins (LCN1_HOMSA), of

rat OBPII (OBP2_RATNO), of bovine lactoglobulin BLG (LACB_BOSTA), of mouse MUP (MUP6_MUSMU), of human RBP (RBP_HOMSA), of bovine OBP (OBP_BOSTA), of rat MUP (MUP_RATNO) and of porcine OBP (OBP_SUSSC).

The residues in the dark gray boxes are identical and those in the light gray boxes are similar. The elements of secondary structure predicted with the DSC program are underlined and the amino acid residues are in italics. The β-sheets and the α-helices are numbered for hOPBIIa and b. The anticipated signal peptide cleavage site is indicated with an arrow (AAA↓LS) at position 15. Nonaligned sequences of very divergent forms of spliced genes, hOPBaδ (OPB2adHOMSA) and hOBPbγ (OBP2bgHOMSA) have been added at the bottom of the alignment after the analysis.

FIGURE 6: RT-PCR analysis

(A) Detection of the LCN1, LCN1b and LCN1c RT-PCR products with their specific probes, and (B) detection of the hOBPIIa and hOBPIIb RT-PCR products with their specific probes. (C) Control G3PDH RNA. The sizes are indicated in bp.

FIGURE 7: Tissue location of the hOBP mRNAs

Sections of the middle meatus (A, B), of the conchae (C, D), of the prostate (E, F), of the vas deferens (G, H), and of the mammary glands (I, J) were hybridized with digoxigenin-11-UTP-labeled hOBP riboprobes. The hybridization with a sense hOBP riboprobe did not reveal any signal (B, D, F, H, J). A specific hybridization signal is obtained with an antisense riboprobe (A, C, E, G, I). The arrows indicate different structures: AC: acinar cells, EC: epithelial cells, GC: glandular cells, SD: secretory duct and L: lumen (X200).

FIGURE 8: Phylogenetic distance tree for vertebrate lipocalins

The usual abbreviations for lipocalins were used. Indicated after the symbol "_" are the first three letters of the genus followed by the first two letters of the species.

FIGURE 9: Scheme of the evolution of the LCN1-OBPII gene subfamily during evolution

The boxes indicate the exons arranged on a line representing the genomic DNA. The "//" between the lines indicate that the loci are not consecutive, without it being possible to determine the order. The large cross under the symbol "?" illustrates a partial duplication or complete duplication event with a subsequent genomic deletion for the LCN1c locus (choice indicated by the symbol "*"). The smallest cross indicates the removal of the seventh exon which appears to be specific for humans because of the numerous repeat sequences Alu present in this genomic region and because the rat has two VEGP genes. The symbol "**" marks the exon recruitment step based on the data present and those of the literature; this step may have appeared at any time after the duplications of the loci and may have been sequential.

FIGURE 10: RT-PCR analysis, in the oral sphere and the genital sphere, of the expression of the other known human lipocalins

Definition of novel OBP and VEGP functions for the RBP (retinol-binding protein) and ApoD (apolipoprotein D) proteins.

FIGURE 11: Study of the polyclonal antibody against the hOBPII proteins

The fusion protein is loaded in wells 1 and 3. A sample of human nasal mucus is loaded in wells 2 and 4. The proteins contained in wells 1 and 2 were revealed with Coomassie blue staining. The antibody hybridized on the membrane obtained by Western blotting specifically reveals the GST-hOBPIIb protein between 30 and 42 kDa (well 3) and the 18 kDa hOBPIIa and hOBPIIb (well 4) in the human nasal mucus.

FIGURE 12: Immunohistochemistry on the tissues of the olfactory apparatus

Strong immunoreactivity can be observed on the section of septum (green staining, A) on the two sections of concha (C and D) and on the section of middle meatus (F). The negative controls prepared for the septum (B), for the concha (E) and for the middle meatus (G) show no reactivity. The cell nuclei are located by DAPI staining (blue staining). The magnifications used are X 100 for the sections A, B, C, E and G, and X 200 for the sections D and F.

FIGURE 13: Immunohistochemistry on tissues of the oral sphere

These two tissues, lacrymal glands (A, B) and von Ebner's glands (E, F) show strong immunoreactivity compared to the results obtained with the pre-immune serum (C for the lacrymal glands and D for the von Ebner's glands). Magnification X 200.

FIGURE 14: Immunohistochemistry on the mammary glands and lung

Strong immunoreactivity is visible on the sections of mammary glands (A and B) compared to the negative

control (C) prepared using pre-immune serum. Strong immunoreactivity is also observed on the sections of lung (E and F), again compared to the negative control (D). Each observation is made at a magnification of X 100 for A and C and at a magnification of X 200 for B, D, E and F.

FIGURE 15: Demonstration of the hOBPII proteins in various nasal structures: concha (A: 200x, B: 400x), septum (C: 300x, D: 600x), middle meatus (E: 400x, F: 600x)

Detection of the hOBPII proteins with a rabbit polyclonal serum revealed by peroxidase activity detected by DAB. AC: secretory acinar cell, P: DAB precipitate (hOBPII detection).

FIGURE 16: Demonstration of the LCN1 protein in various nasal structures: concha (A: 200x, B: 500x), septum (C: 100x, D: 500x), middle meatus (E: 200x, F: 600x)

Detection of the LCN1 protein with a rabbit polyclonal serum revealed by peroxidase activity detected by DAB. AC: secretory acinar cells, P: DAB precipitate (LCN1 detection).

FIGURE 17: Demonstration of the ApoD protein in various nasal structures: concha (A: 500x, B: 700x), septum (C: 300x), middle meatus (D: 400x)

Detection of the ApoD protein with a mouse monoclonal antibody revealed by peroxidase activity detected by DAB. AC: secretory acinar cell, P: DAB precipitate (ApoD detection).

FIGURE 18: Demonstration of the hOBPII protein in various structures of the oral sphere: lacrymal gland (A: 200x, B: 600x), von Ebner's glands (C: 100x, D: 400x), lung (E: 200x, 800x)

Detection of the hOBPII protein with a rabbit polyclonal serum revealed by peroxidase activity detected by DAB. AC: secretory acinar cell, P: DAB precipitate (hOBPII detection).

FIGURE 19: Demonstration of the hOBPII proteins in various placental structures: chorionic villi (A: 300x, B: 800x), amniotic membrane (C: 800x), cord (D: 600x)

Detection of the hOBPII proteins with a rabbit polyclonal serum revealed by peroxidase activity detected by DAB. L: lacuna (contains maternal blood), C: fetal capillary, Am: amniotic membrane, E: epithelial cells derived from the extra-embryonic ectoblast, MB: basement membrane, M: mesenchymal layer, P: DAB precipitate (hOBPII detection).

FIGURE 20: Demonstration of the LCN1 protein in various placental structures: chorionic villi (A: 200x, B: 800x), amniotic membrane (C: 600x), cord (D: 500x)

Detection of the LCN1 protein with a rabbit polyclonal serum revealed by peroxidase activity detected by DAB. L: lacuna (contains maternal blood), C: fetal capillary, Am: amniotic membrane, E: epithelial cells derived from the extra-embryonic ectoblast, MB: basement membrane, M: mesenchymal layer, P: DAB precipitate (LCN1 detection).

FIGURE 21: Demonstration of the ApoD protein in various placental structures: chorionic villi (A: 100x, B: 500x), amniotic membrane (C: 400x), cord (D: 400x)

Detection of the ApoD protein with a mouse monoclonal antibody revealed by peroxidase activity detected by DAB. L: lacuna (contains maternal blood), C: fetal capillary, Am: amniotic membrane, E: epithelial cells derived from the extra-embryonic ectoblast, MB: basement membrane, M: mesenchymal layer, P: DAB precipitate (ApoD detection).

EXAMPLES

EXAMPLE 1: MATERIALS AND METHODS

1.A. Human chromosome 9 cosmid library

A copy of the library of cosmids specific for human chromosome 9, LL09NC01P, constructed by Dr. J. Allmeman (Biochemical Sciences Division, Lawrence Livermore National Library, Livermore, CA USA) under the aegis of the National Gene Library Project, supported financially by the American Department of Energy, was used. The library was screened and the clones analyzed as previously described (Lacazette *et al.*, 1997).

1.B. Cloning and sequence analysis

A Lambda gt11 human testes cDNA library (Clontech) (10^7 p.f.u.) was amplified by 30 polymerase chain reaction (PCR) cycles (94°C 45 sec, 54°C 45 sec, 72°C 1 min 30 sec) with the primer oliEST58 CCTGCAGGTACATGAGCTTCC and 5' or 3' screening amplimers for inserts located on the arms of the lambda gt11 vectors. A nested PCR was then carried out with oliEST26 CGCTGTATTTGCCAGGCTCC and oligonucleotides specific for the arm of the vector. The PCR products were subcloned into the pGEM-T(r)

vector, which made it possible to obtain the 5' end of the hOBPII gene cDNAs.

The sequences obtained using the standard oligonucleotide pGEM-T(r) and a dye terminator-based sequencing ready reaction mix (Applied Biosystems) were separated by electrophoresis using an ABI PRISM 377 automatic sequencer (Perkin Elmer) and were then analyzed with the sequence nagivator 1.0.1. software (Perkin Elmer). Full length cDNA clones of hOBPIIa (hOBPIIaα, hOBPIIaβ, hOBPIIaδ, hOBPIIaγ) and hOBPIIb (hOBPIIbα, hOBPIIbβ, hOBPIIbγ) were obtained from the RT-PCR by purifying the bands of interest according to the manufacturer's instructions (Qiagen gel extraction kit) or, for the weakly expressed alternative forms, by subcloning the nested PCR products into a pGEM-T(r) vector.

1.C. RT-PCR analysis

Tissue samples were collected from 45 to 55 year old Caucasian individuals, in agreement with the French regulations in force. The total RNA is extracted according to a single-step method using the RNA NOW® reagent, according to the manufacturer's (Biogentex) instructions. 5 µg of total RNA were reverse-transcribed in a final volume of 20 µl using 0.5 ng of oligonucleotide GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT with the Superscript® preamplification system (Gibco BRL). Three µl of these reactions were then used for the following PCRs. The expression of the specific mRNAs was determined by PCR using: the primers TL: CCTCTCCCAGCCCCAGCAAG and AP: GACTCGAGTCGACATCG for the LCN1-type genes (LCN1, LCN1b, LCN1c) and, for the hOBPII-type genes, the primers DE: CGCCCAGTGACCTGCCGAGGTC, and FI: CTTTATTTGGAGTCAGGTGGGTG. As a quality control for the RNAs the primers G3PDH1: CTCTGCCCCCTCTGCTGATG and G3PDH2: CCTGCTTCACCACCTTCTTG for the G3PDH gene were used; the G3PDH gene is considered to be constitutively

expressed in all cell types. 32 PCR cycles (94°C 45 sec, 54°C 45 sec, 72°C 2 min 30 sec) were performed and the amplification products were separated on 1% agarose gel. The DNA was transferred onto a Hybond N+® membrane.

For the detection of the expression of the various genes, several oligonucleotides specific for the respective genes were synthesized:

- olLCN1: GACTCAGACTCCGGAGATGA,
- olLCN1b: AACTCAGACACCAGAGATGA,
- olLCN1c: GACTCAGATCCCGGAGATGA,
- ol5: CCAGGAGGGACCACTACA specific for the hOBPIIb gene,
- ol4: CCGGGACGGACGACTACG specific for the hOBPIIa gene,
- G3PDH3: CTCATGACCACAGTCCATGC.

The oligonucleotides are labeled with γ^{32}P-ATP using T4 kinase (Applied Biosystem); the labeled oligonucleotide hybridizations are performed at 42°C. The final wash is performed in a solution of 0.1 X SSC, 0.1% SDS at 48°C for 20 min. The specificity of the oligonucleotide hybridization reactions is controlled using samples of digested cosmid DNA (p233G2 for LCN1 and hOBPIIa, P19E7 for LCN1b and hOBPIIb, and P181A9 for LCN1c) loaded onto the gel with RT-PCR products.

1.D. <u>Genotyping and linkage analysis</u>

The genotyping is carried out with PCR reactions using 100 ng of genomic DNA originating from 8 CEPH reference families, and using the oligonucleotides oli9 TGTTCGGGAACGCAGCTT and oli10a TGCCGCTGTCCCCACGTCGG. The thermocycler parameters consist of an initial cycle at 94°C for 10 min followed by 30 cycles at 94°C for 30 sec, 55°C for 30 sec and 70°C for 45 sec, and then a final elongation step of 10 min at 70°C. The PCR products are then analyzed on a 3% agarose gel. The information regarding the markers for chromosome 9 may

be obtained at the following internet address (hTTP://galton.ucl.ac.uk); the analyses were carried out using the linkage study tools previously described in Lacazette *et al.* (1997). The haplotypes are reconstructed manually according to the recombination events previously described in family 1362 (Attwood *et al.*, 1994).

1.E. Secondary structure predictions

A multiple alignment of the lipocalin proteins for which the crystallographic structures have already been described (Monaco *et al.*, 1992; Spinelli *et al.*, 1998), with the hOBPIIa and hOBPIIb proteins, was obtained using the Clustalx software (ftp.infobiogen.fr). The putative secondary structures were determined with the DSC (Discrimination of protein Secondary structure Class) program developed by R.D. King and M.J.E. Sternberg (http://bioweb.pasteur.fr/seqanal/interfaces/dsc-simple.html). The secondary structures of the proteins corresponding to the alternatively spliced forms are presumed to be identical to the conventional forms before the reading frame shift; after the shift, the structural prediction is made with a single sequence and is performed using the Predator software (http://pbil.ibcp.fr/cgi-bin).

1.F. Phylogenetic analysis

The lipocalin proteins, the sequences of which are entirely known, were aligned three times consecutively using the Clustalx software (ftp://ftp.infobiogen.fr). Distances in the phylogenetic tree were calculated with Clustalx and plotted with Njplot.

1.G. *In situ* hybridization

Serial sections cut on a cryostat (8 µm thick) are collected on SuperFrost® Plus slides (Menzel Glazer)

and stored at -80°C. Antisense and sense RNA probes are obtained according to standard techniques using T7 or SP6 polymerase, from a matrix obtained by digestion of the phOBPIIaP2 cDNA clone with NcoI or PstI restriction enzymes, and using DIG-11-UTP (Boehringer Mannheim) (the probe is approximately 150 nucleotides long). The matrices digested with PstI and transcribed with T7 RNA polymerase correspond to the antisense probe and the matrices digested with NcoI and transcribed with the SP6 RNA polymerase correspond to the sense probe. Tissue sections are fixed in 4% paraformaldehyde for 15 min and then rinsed for 5 min in 2 X PBS. After acetylation (2 × 5 min in a triethanolamine (TEA) buffer, pH 8, containing 0.25%, v/v, of acetic anhydride), the tissue sections are prehybridized at 60°C for 15 min in 50% formamide/1 X SSC. The labeled probes are applied to each of the sections in 50 µl of hybridization buffer (50% formamide, 1 X Denhardt's, 500 µg/ml total tRNA, 10% Dextran sulfate, 10 mM dithiothreitol). The sections are covered and then incubated in humid chambers at 50°C overnight. After hybridization, the slides are immersed at 55°C in a washing buffer (50% formamide, 1 X SSC) for 2 hours. Each slide is then rinsed twice 5 min in 2 X SSC at ambient temperature, and then treated for 30 min with 10 mg/ml of RNAse at 37°C and, finally, immersed in a washing solution (50% formamide, 2 X SSC) for 2 hours at 55°C. The slides are then placed in 0.1 X SSC for 15 min at 55°C. The immunodetection is performed using an alkaline phosphatase-conjugated anti-DIG antibody (Fab fragments) according to the Boehringer Mannheim protocol. The sections are examined at different magnifications using an Axiophot microscope (Zeiss).

1.H. <u>Production of a monoclonal antibody against the $hOPB_{II}$ proteins</u>

Two hundred and fifty µl of fusion proteins were given to the company Agro-Bio for the production of

antibodies. Rabbits (New Zealand White SPF) are injected on days 0, 14, 28 and 42. The sera are taken on days 0, 35, 49 and 63. In order to test the antibody, the fusion protein is loaded in lanes 1 and 3 (figure 11), while human nasal mucus is loaded in lanes 2 and 4. Lanes 1 and 2, and also the lane corresponding to the molecular weight ladder, correspond to revelation with Coomassie blue. Lanes 3 and 4 correspond to detection of the proteins recognized by the antibody using the Western blot technique. Lanes 1 and 3 show the presence of a series of truncated recombinant proteins between 30 and 42 kDa derived from a relatively inefficient synthesis due to the presence of many codons which are rare in the bacterium, within the hOBPIIb sequence. However, this production of recombinant protein was sufficient to produce a polyclonal antibody of good quality, as indicated by the revelation, in lane 4 corresponding to the nasal mucus, of an 18 kDa band specific for the $hOBP_{II}$ proteins.

1.I. <u>Immunohistochemistry</u>

1.I.1. 8 µm serial sections of various tissues are fixed with paraformaldehyde (5 min) and rinsed three times with 1 X PBS (15 min) then incubated for 30 min in a 3% solution of BSA in 1 X PBS, before being incubated overnight in the presence of the anti-fusion protein antibody in a 1 X PBS solution. After three rinses in 1 X PBS (15 min), a rabbit anti-IgG antibody coupled to FITC (green staining by fluorescence) in a 1 X PBS solution is placed in contact with the slides for 3 hours. After rinsing in 1 X PBS, a solution of DAPI (100 ng/µl in 1 X PBS) is applied for 10 min (counterstaining of the cell nuclei in blue). After three rinses in 1 X PBS (15 min) the slides are mounted in water-glycerol (50/50). The analysis is carried out in the presence of DAPI and FITC filters with the aid of a CDD camera using an integration time of between 4 and 32 ms.

1.I.2. The tissues are frozen at -80°C in the presence of OCT. 7 micrometer serial sections are cut using a microtome. The tissues are then fixed with 4% paraformaldehyde in PBS. The slides are incubated for 10 min in a solution of methanol in the presence of 4.5% of H_2O_2, and then rinsed three times in PBS for 5 min. After incubation for 30 min in a 3% bovine serum albumin solution, the first antibody (diluted to 1/350) is incubated overnight at ambient temperature (rabbit polyclonal antibodies for hOBPII and LCN1, and mouse monoclonal antibody for ApoD). Then, after three 15 min rinses in PBS, the second antibody (diluted to 1/300) (anti-rabbit for hOBPII and LCN1, and anti-mouse for ApoD) is incubated for 3 h at ambient temperature. After three rinses in PBS for 15 min, peroxidase revelation with DAB is carried out for 2 to 10 min, using the "Vector" kit. The slides are then rinsed with water and counterstained with Mayer's hematoxylin for 2 min. After two rinses with water, the slides are dehydrated by passing them successively through baths containing 70°, 95°, 100° and 100° ethanol solutions and, finally, two baths of toluene. The slides are then mounted in Canada balsam and observed using a Zeiss microscope.

EXAMPLE 2: IDENTIFICATION OF A GENE HOMOLOGOUS TO LCN1 LOCATED ON HUMAN CHROMOSOME 9

The identification of the LCN1 complementary DNA encoding the lipocalin from human tears (Lassagne and Gachon, 1993) has previously been reported, as has the location of the gene on chromosome 9q34 (Lassagne *et al*, 1993; Lacazette *et al*, 1997). The two genes encoding the von Ebner's gland proteins 1 and 2 (Kock *et al*, 1994) correspond to the proteins homologous to LCN1 in rats; this poses the question of whether other genes encoding an LCN1 protein exist in the human genome. *In situ* hybridization experiments (Lassagne *et*

al, 1993) and also analyses with somatic hybrids (Lassagne *et al*, 1995) indicate that, if they exist, these additional human genes must be located in chromosomal region 9q34.

A library of cosmids specific for human chromosome 9 (LL09NP01) was screened with the human LCN1 cDNA probe; 26 cosmids were identified; they were digested with EcoRI or PvuII, and then hybridized successively with the LCN1 DNA and various oligonucleotides (figure 1).

The cosmids are divided into 3 groups. The first group (clones P32H3, P41B5, P63B6, P92H20, P109C6, P145H6, P195B4, P233G2, P233F2, P265D4 and P276H8) correspond to the cosmids containing a sequence of the previously isolated LCN1 gene (accession number: L14927) made up of 7 exons (Holzfeind and Redl, 1994). The second group (clone P19E4, P19E7, P42H9, P98H5 and P142H8) corresponds to the sequence LCN1b homologous to LCN1 (clone P19E4) (accession number: Y10826) from the promoter up to the 6th exon, which then diverges. A third group of cosmids (clone P110C1, P174E4, P174E5, P181A9, P181B6, P211A7, P238C6 and P291E1) contains a region LCN1c, established from the partial sequencing of the clone P181A9 (accession number Y10827), which is highly homologous to LCN1 only from the promoter up to exon 2. Thus, LCN1 is the only gene which contains the 7th and last exon. In addition, the TATA boxes are degenerate in the promoters of the LCN1b and LCN1c genes.

EXAMPLE 3: IDENTIFICATION OF GENOMIC DUPLICATIONS CONTAINING THE LIPOCALIN GENES AND MAPPING OF CHROMOSOME REGION 9Q34

During the period in time over which the LCN1 gene family was identified, a vast project of physical mapping led to the identification of cosmid contigs partially covering chromosome region 9q34 (Nahmias *et al*, 1994, Van Slegtenhorst *et al.*, 1995, Hornigold *et*

al, 1997); the sequencing of the corresponding clones was performed by a group from the Massachusetts Institute of Technology (Boston, USA).

Analyses comparing sequences in the database, for LCN1, LCN1b and LCN1c, revealed strong homologies between the previously described LCN1 gene (accession number L14927) and the three cosmid sequences: P161A1 having the accession number AC002098, P203H12 with the accession number AC000396 and P161G2 with the accession number AC002106.

A very fine analysis, in particular of the 3' region of the LCN1 genes, revealed that the LCN1c sequence corresponded to the clone having the accession number AC002106, LCN1b to the number AC002098 and LCN1 to the number AC000396; there is, however, an insertion of 60 base pairs at position 12360 of the clone having the accession number AC000396 compared to that having the number L14927. Additional analyses revealed that these sequences are homologous over a region which is much larger than the LCN1 genes. As illustrated by the dot plot analyses (figure 2), these three cosmids correspond to duplicated genomic regions; the cosmids P161A1 (AC002098) and P203H12 (AC000396) are homologous over their entire sequences; the cosmid P161G2 (AC002106) is homologous to the sequence only for sequences located upstream of intron 3 of the LCN1 gene.

The relative positions of these duplicated regions of human chromosome 9 were analyzed. The results were compared with the physical map established in the context of the physical mapping project by restriction analyses of the cosmids (Hornigold *et al*, 1997) using the same library LLN09NC01P. Approximately the same cosmids were found in the two searches for the LCN1c locus; the sequence comparison for the cosmid P181A9 reveals that its T3 end contains the Surf 5 gene (figure 1). In addition, AC002106 is located in a

sequence contig between ABO and the Surfeit locus, confirming the location of LCN1c. P161A1 and P203H12 have less precise locations using the data from Hornigold and from his collaborators because of the limit of the restriction mapping strategy for the duplicated regions. P203H12 (AC000396) appears to correspond to the LCN1 gene which was previously located close to the D9S1826 marker (Lacazette *et al.*, 1997), with the exception of a 60 base pair insertion.

In order to test whether P203H12 may correspond to a 4th duplicated region, the region was tested by PCR on genomic DNA from 150 unrelated individuals using primers specific for AC000396 and L14927. A single band corresponding to the length of the AC000396 sequence was detected (data not shown), thus proving that P203H12 contains the previously described LCN1 gene. The LCN1b locus was not positioned precisely in the region, as demonstrated by the fact that the AC002098 ends and also the ends of the cosmids containing LCN1b detect no homologous sequence in the databanks.

The sequence analysis made it possible to identify a putative minisatellite at position 3177-3724 of the cosmid corresponding to AC002098 (figure 1); a rare polymorphism was thus demonstrated in a population of 20 unrelated individuals (PIC = 0.05). This new polymorphic marker is informative only in family 1362 of the 8 CEPH reference families tested.

The linkage analysis revealed two-point lod scores greater than 3 at theta = 0 for the D9S275 and D9S1818 markers. The reconstruction of the haplotype confirmed the location of the LCN1b gene between the D9S1811 and D9S67 markers in chromosome 9q34 (figure 1).

EXAMPLE 4: IDENTIFICATION OF TWO NEW LIPOCALIN GENES

The sequence analyses revealed new data.

The comparison of AC002098 to the databank revealed similarities between this sequence and the lipocalin genes. In addition to the region 21000 to 27000, which contains the LCN1b gene, the region around position 2150 contains sequences which exhibit similarities with the sequences encoding the rat type II odorant-binding proteins, lipocalins from tears and also the EST AA460385.

In parallel, the comparison of the AC000396 sequence to the databank revealed, in addition to the region 11100 to 17100 containing the LCN1 gene, similarities for the same group of sequences in the region 36600 to 37800.

The EST AA460385 expressed in human testes corresponds to 4 exons of the new lipocalin gene present in the clone P161A1 (AC002098). A putative 50 base pair exon similar to the EST sequence is also present at the end of the clone P203H12 (AC000396).

The inventors have therefore concluded that a new lipocalin gene exists at a position distal to LCN1b. The inventors have put forward the hypothesis that, subsequent to a genomic duplication, a second gene orthologous to the gene encoding the EST is present in the 3' region of LCN1.

The sequencing of the cosmid P233G2 containing this region (figure 1) with oligonucleotides defined from the EST in fact revealed the presence of another lipocalin gene at a locus 20 kb distal to the LCN1 gene. In order to identify the first exons of the two new genes, nested PCRs on cDNA clones originating from a testes library were carried out between the oligonucleotides located in the 5' region of the EST and the arms of the vector. The PCR products were cloned and their sequence revealed three additional exons going by the genomic sequence (figure 1). A TATA box is present upstream of

the first exon in both cases (figure 3). The two mRNAs, hOBPIIa corresponding to the gene located downstream of LCN1 and hOBPIIb located downstream of LCN1b (figures 3 and 4), are 97.5% and 63% identical to those corresponding to the LCN1 gene. The intron/exon organizations of these two genes are consistent with the lipocalin family.

In addition, the deduced protein sequences (170 amino acids) confirm that they belong to the lipocalin family. The two putative mature proteins hOBPIIa (MW = 17.8 kDa) and hOBPIIb (MW = 18.0 kDa) are 89% identical. These two novel proteins contain a putative signal peptide of 15 amino acids (figure 5), the conserved residues GXWY at position 27 to 30, and two cysteines which may be involved in a disulfide bridge (position 74 to 166).

However, the paired sequence comparisons with the other lipocalins indicate weak conservation of the amino acid residues (15 to 25%), with the exception of human TL/VEG encoded by the LCN1 gene (mean value 43% over 155 amino acids) and rat OBPII (mean value 45.5%). Moreover, the calculated isoelectric points of hOBPIIa and of hOBPIIb are 7.85 and 8.72, respectively, whereas those of the lipocalins are acid (in the region of 4.5), with the exception of rat OBPII (PI = 9.01). The prediction for the secondary structures of the hOBPIIa and hOBPIIb proteins with the DSC program, by carrying out multiple alignments with the sequences of lipocalins of known structures, indicates the presence of 8 strands of antiparallel β-sheets which may allow the formation of a calyx, followed by an α-helix and a final β-sheet, in agreement with the known data for the structures of the lipocalins (figure 5).

EXAMPLE 5: STUDY OF THE EXPRESSION OF THE LIPOCALIN GENES IDENTIFIED

In order to specify whether LCN1b, LCN1c, hOBPIIa and hOBPIIb are expressed genes, RT-PCR analyses were carried out on 18 different human tissues known to produce lipocalins (figure 6).

Two pairs of primers were synthesized; one recognizes the LCN1-type mRNAs and the other the hOBPII-type mRNAs. The hybridization of the oligonucleotides specific for the genes, on membranes on which the RT-PCR products are immobilized after transfer via the Southern technique, and then the subcloning of the hybridized products revealed that LCN1b and LCN1c are not expressed in any of the products tested, whereas the LCN1 mRNAs are detected in the lacrymal gland, the sweat gland, the von Ebner's glands, the nasal septum, the epithelium of the nasal concha, as well as the placenta and mammary glands (figure 6A). These data and the sequence information, which indicates that the TATA box and the last exon have been lost, make it possible to affirm that LCN1b and LCN1c are pseudogenes which do not contribute to the formation of the human LCN1 proteins.

On the other hand, the hOBPIIa and hOBPIIb genes are expressed, which confirms the previous detection thereof in the cDNA libraries. Surprisingly, although both proteins hOBPIIa and hOBPIIb are very similar over their entire sequence, including in the 1.5 kb promoter region, their expression profiles are different (figure 6B). The hOBPIIa protein is strongly expressed in the nasal septum, the middle meatus, the nasal concha, the testes and the placenta, and more weakly in the mammary glands, the lacrymal glands, the sweat glands, the von Ebner's glands and the lung. On the other hand, the hOBPIIb protein is expressed predominantly in the prostate, the testes and the

mammary glands, and more weakly in the submaxillary glands, the nasal septum and the middle meatus.

In addition, the RT-PCR analyses revealed the existence of alternative splicing of the product of transcription of the hOBPIIa gene and of the hOBPIIb gene, which generates, respectively, four and three mRNAs (figures 3, 4, 5 and 6).

The transcription of the hOBPIIa gene generates at least four mRNAs which encode four different proteins; the first mRNA encodes the hOBPIIaα protein which corresponds to the hOBPIIa protein described above. In the hOBPIIa gene, three different splice acceptor sites have been identified for exon 5 (figures 3 and 4), thus forming two other splice variants. A first splice variant has an acceptor site for exon 5 located 49 bp before the previous one (exon 5b); this generates a 725 nucleotide mRNA which encodes the 146 amino acid protein hOBPIIaβ. This protein is identical up to the 8th putative β-sheet and then different with only 16 additional amino acids. A second splice variant has an acceptor site for exon 5 located 65 bp before the previous one (exon 5c); this generates a 741 nucleotide mRNA which encodes a 228 amino acid protein hOBPIIaγ. This protein has the first eight putative β-sheets identical to those of hOBPIIaα, and is then different in the C-terminal region (figure 5) due to a reading frame shift generated by this alternative splicing event; the structure of this C-terminal region predicted by the Predator program is a long bent region containing a 9th β-sheet.

In the case of the hOBPIIb gene, in addition to the hOBPIIbα mRNA previously described, a 106 bp super-numerary exon (exon 3b) between the previous exons 3 and 4 was identified (figure 3). This longer mRNA (782 nucleotides) encodes a 165 amino acid protein hOBPIIbβ. From a protein structure point of view,

hOBPIIbβ is identical to hOBPIIbα up to the 5th putative β-sheet and then differs due to a reading frame shift. The predictions from the computer programs indicate that the motif ALWEALAIDTRLK is an α-helix which is just behind the fifth β-sheet. Two additional β-sheets may be present in the long C-terminal portion.

None of the alternative splicing variants for one gene is detected symmetrically for the other gene, although the putative splice acceptor and donor sites may be present (figure 3). In addition to these protein motifs which conserve the conventional structure of a lipocalin, we have demonstrated, for the two genes hOBPIIa and hOBPIIb, the existence of a small amount of mRNA which has undergone alternative splicing and encodes proteins, the sequence of which is not directly related to the lipocalins. mRNAs encoding hOBPIIaδ and hOBPIIbγ, which lack the sequence encoding exon 2 and which have, respectively, exon 5b and exon 5, encode putative secreted proteins of 147 and 85 amino acids, respectively (figures 4 and 5); these proteins diverge from the previous proteins from the 24th amino acid.

In order to identify the cells expressing the hOBPII genes, the inventors hybridized digoxigenin-labeled sense and antisense ribonucleic acid probes on tissue sections (figure 7). The hOBPII mRNAs are detected in the acinar cells of the middle meatus and of the nasal conchae, and also epithelial cells of the conchae; this supports the idea that the hOBPII proteins are involved in olfactory function. In addition to the production in the oral sphere, mRNAs encoding hOBPII were detected in the genital sphere, in particular in the glandular cells of the prostate, and in the secretory epithelial cells of the vas deferens. No signal was detected in the male gonads, which suggests that the expression of the hOBPII genes detected in the RT-PCR experiments corresponded to the presence of additional ducts in the tissue preparation (rete testis and lobuli

epididymidis). In combining these results from the detection of all of the mRNAs encoding hOBPII with those of the RT-PCR approach, it appeared that five hOBPII proteins (hOBPIIaα, hOBPIIaβ, hOBPIIaγ, hOBPIIbα, hOBPIIbβ) are secreted by the epithelial cells of the ducts of the male gonads, and also acinar cells of the middle meatus and of the nasal conchae; in these cells, the mRNAs encoding hOBPIIa are highly predominant. Only the two hOBPIIb proteins (hOBPIIbα and hOBPIIbβ) are secreted by the epithelio-glandular cells of the prostate and of the mammary glands.

EXAMPLE 6: PHYLOGENETIC ANALYSIS AND CLASSIFICATION OF THE LIPOCALINS

The dotplot analysis (figure 2) and also the genomic duplications which we revealed indicate a common origin for the LCN1 and hOBPII genes. To clarify the relationships between the members of the lipocalin family and to verify that we identified the human gene orthologous to rat OBPII (Dear *et al.*, 1991), we constructed a phylogenetic distance tree with the vertebrate lipocalins (figure 8).

The inventors demonstrated nine main groups of lipocalins derived from a common precursor:

- the apolipoprotein and retinol binding protein (RBP) family (group 1);
- the prostaglandin D synthase and neutrophil gelatinase-associated lipocalin precursor group (group 2);
- the alpha-1-microglobulin/bikunin (protein HC) subfamily (group 3);
- the orosomucoid subfamily (A1AG, A1AH, A1AI) (group 4);
- the oral sphere subfamily 1 (OBPII-type-LCN1/VEGP, VNSP I and II, LALP, CanF1) (group 5);
- the lactoglobulin subfamily (group 6);

- the subfamily of proteins secreted by the lizard epididymis (group 7);
- the oral sphere subfamily 2 (mouse major urinary proteins (MUP)) (group 8);
- the oral sphere subfamily 3 (OBP1, mouse OBPII, aphrodisin, probasin, BD20) (group 9).

The present invention relates more particularly to groups 5, 8 and 9.

Group 5 contains the hOBPII proteins which are closely related from an evolutionary point of view to rat OBPII. Taken together, the present results indicate that the human hOBPII genes are orthologs of the rat OBPII gene. This group also contains the LCN1-VEGP proteins of various species. Taking into consideration the genomic organization of the hOBPII-LCN1 genes, the data from the tree illustrate the duplication event (arrow) which gives rise to the ancestral hOBPII and LCN1 genes from their common lipocalin precursor (figure 9). More recently, the original duplications of the 50 kb region containing hOBPII-LCN1 (arrow) generated, in humans, the two hOBPII genes and the LCN1 gene and its pseudogene LCN1b. The additional duplication which gave rise to the LCN1c pseudogene is partial in the human genome and does not produce a functional protein, and is therefore absent from the present tree. In addition, the two rat VEG proteins are more closely related to one another in the phylogenetic tree than to the human LCN1 protein. The situation is identical for the two human OBPII proteins compared to the rat OBPII protein. These results are in favor of a process of gene conversion for at least some lipocalin genes. This may also be correlated to the fact that these genes are located on the same chromosomal arm.

Group 9 corresponds to the oral sphere family 3 and contains OBPs which have already been described. The aphrodisin protein has already been described as a

pheromone transporter (Henzel *et al.*, 1988) and appears to be orthologous to the rat and mouse OBP1 protein, with two paralogous OBP1 genes. Finally, the olfactory lipocalins produced by the Bowman's glands of *rana pipiens* (OLFA RANPI), which are considered to be potential odorant transporters in frog mucus, are not related to any putative OBP group (groups 5, 8 and 9), thus suggesting the existence of other OBP categories.

Some lipocalins have been described as being allergens. The major allergen of dogs, the CanF1 protein, mainly expressed in the von Ebner's glands (Konieczny *et al.*, 1997), appears in the tree to be the dog VEG protein. This result, which shows that LCN1 may be allergenic, prompts the inventors to propose that the OBPII proteins are involved in allergic processes. Similarly, the major allergen of horses, EquC1, expressed in the liver and the salivary glands (Grégoire *et al.*, 1996), appears to be orthologous to an MUP protein member (group 3). The major allergen of cows (BD20 BOSTA), which is present in the oral sphere family 2 (group 4) is more closely related to probasin. CanF2, which is mainly expressed in the parotid gland (Konieczny *et al.*, 1997), does not appear to be orthologous to a previously described lipocalin.

The present invention has made it possible to reveal the existence of a genomic duplication at the q34 locus of human chromosome 9, which harbors a gene family of the LCN1 type; this family comprises, besides the LCN1 gene, which has been previously described, two pseudogenes and also two hOBPII genes which are paralogous to LCN1. The inventors have revealed that the hOBPII-LCN1 family results from consecutive genomic duplication events. The sequences and also the genomic organization revealed that the LCN1 and hOBPII genes derive from a common ancestor and were generated by means of tandem duplication. The sequence comparisons demonstrated that the hOBPIIaα and hOBPIIbα proteins are the hOBPII

protein forms closest to LCN1; this is corroborated by the fact that the size of exon 5 of LCN1 is 121 bp, which corresponds to the size of exon 5a of the hOBPII protein, and by the fact that no exon 3b was found in the LCN1 protein. Similar results are obtained when this comparison is carried out with the other seven exons of the lipocalins. These data are in favor of the diversity of the hOBPII proteins being acquired through the integration of surrounding additional genomic DNA at the upstream splice acceptor site for exon 5 of hOBPIIa or through the recruitment of a supernumerary exon (exon 3b) for the hOBPIIb protein. The hypothesis of a recruitment of exons rather than a decrease in the size of the exons or than a loss of exons is more probable.

The inventors showed that the hOBPII and LCN1 genes encode proteins involved in various functions, as shown by the expression of these genes in both the oral sphere and the genital sphere. In addition, the inventors showed, by the phylogenetic analysis (example 6), that several different proteins could contribute to the same odorant-binding protein function; thus, three lipocalin subfamilies (groups 5, 8 and 9) corresponding to several proteins are found expressed in the oral sphere, and in particular in the nasal and buccal glands.

This analysis, showing that several lipocalins contribute to the same physiological function and that the same protein could contribute to various functions, led the inventors to analyze the expression of the other human lipocalins in all of the tissues studied (figure 10). This led to the demonstration that the gene encoding apolipoprotein D is expressed in the glands of the nasal concha and of the middle meatus, thus making it a potential odorant-binding protein. Similarly, retinol binding protein (RBP) is expressed

by the nasal epithelium and may also contribute to this function.

The inventors therefore propose the inclusion of the ApoD and RBP proteins in the human OBP family and include them in all the claims relating to these human OBPs.

The inventors have demonstrated that the transcription of the two hOBPII genes generates many alternative transcripts which encode different proteins, the structure of which is compatible with that of a hydrophobic ligand transporter.

The inventors have demonstrated the expression of the two hOBPII genes in the oral sphere (nasal glands, von Ebner's glands, submaxillary glands, lacrymal glands, lung). The inventors have also demonstrated that the hOBPII proteins according to the invention are produced by the cells of the genital sphere; the hOBPII gene is mainly expressed in the prostate, the vas deferens and the mammary glands, while the expression of the hOBPIIa gene is restricted to the vas deferens. The in situ analyses have revealed that the mRNA is produced by the glandular cells of the prostate and epithelial secretory cells of the vas deferens, supporting the idea that the corresponding proteins are secreted into the seminal fluid. The proteins according to the invention may therefore be involved in the reproductive function, but also in all the other functions usually attributed to lipocalins.

EXAMPLE 7: PRODUCTION OF A RECOMBINANT HOBPIIBα PROTEIN IN A PROKARYOTIC SYSTEM

A PCR on the plasmid DNA of an hOBPIIbα clone using the primers BIIa/b (5' GTC GGA TCC CTG TCC TTC ACC CTG GAG G 3'), a sense oligonucleotide beginning 45 bases after the protein-initiating ATG, and XIIb (5' GTC CTC GAG

GTG TTC GGG AAC GCA GCT TC 3'), an antisense oligonucleotide preceding the stop codon of the hOBPIIbα protein, made it possible to amplify all of the DNA encoding the secreted hOBPIIbα protein. The BamH I and Xho I enzymatic restriction sites located at the ends of the two oligonucleotides (bases underlined) were used for directional cloning into a plasmid expression vector pGEX-6P1, followed by transformation by electroporation (1800 V, 200 Ω, 25 μF) into a bacterial strain BL21. Synthesis of the recombinant protein is obtained by adding IPTG, at 5 mM final for 3 h, to 250 ml of culture of the strain in LB medium containing 100 μg/ml of ampicillin pre-incubated at 37°C for 2 h. The centrifuged cultures are taken up in 25 ml of TENGN buffer. The lysate is sonicated and recentrifuged. The fusion protein is then purified using 4 ml of beads bearing covalently attached insoluble glutathione (Sigma) per 25 ml of supernatant. After incubation for 4 h, they are washed with 3 volumes of 1M NaCl and then with 10 volumes of 1 X PBS. The elution is obtained by bringing the beads into contact with a glutathione solution (10 mM reduced glutathione, 50 mM Tris-HCl, pH 8.0) for 10 min.

The amount of recombinant protein produced is estimated by migration on polyacrylamide gel and staining with Coomassie blue (figure 11). The specificity of induction of a recombinant protein is tested by Western blot using a goat anti-GST antibody (figure 11), revealed with a peroxidase-coupled goat anti-IgG antibody, using an ECL+plus kit (Amersham). The separation of the fusion protein into two proteins is obtained by proteolysis of 100 μg of pre-dialyzed recombinant protein with 2U of preScission™ protease (Pharmacia Biotech) in 10 μl final containing 1 X enzyme buffer, for 4 h at 5°C.

EXAMPLE 8: DETECTION OF THE hOBPIIs IN VARIOUS TISSUES BY IMMUNOHISTOCHEMISTRY

The location of the $hOBP_{II}$ proteins in the nasal structures (figures 12 and 15), the buccal structures and the lacrymal glands (figures 13 and 18), the mammary glands and the lungs (figures 14 and 18) is revealed by immunohistochemistry.

EXAMPLE 9: DETECTION OF THE hOBPIIs, LCN1 and ApoD IN THE NASAL STRUCTURES BY RT-PCR AND BY IMMUNOHISTOCHEMISTRY

The RT-PCRs carried out using the sets of primers for the hOBPIIs and for LCN1 reveal the presence of the corresponding mRNAs in at least one of the nasal structures. The techniques of immunohistochemical detection both by fluorescence detection and by revelation using peroxidase and DAB reveal the presence of the hOBPII (figure 15) and LCN1 (figure 16) proteins in the epithelial cells of the acini of the nasal glands, and also in the lumina of the secretory ducts. Some can also be detected on mucus attached to the tissues.

The RT-PCRs carried out using the sets of primers for apolipoprotein D reveal the presence of the corresponding mRNAs in at least one of the nasal structures. The techniques of immunohistochemical detection by revelation using peroxidase and DAB reveal the presence of the ApoD protein (figure 17) in the epithelial cells of the acini of the nasal glands, and also in the lumina of the secretory ducts, although the expression appears to be weaker and in fewer acini than for the hOBPIIs.

EXAMPLE 10: DETECTION OF THE hOBPIIs, LCN1 and ApoD IN THE PLACENTA BY RT-PCR AND BY IMMUNOHISTO-CHEMISTRY

The RT-PCRs carried out using the sets of primers for the hOBPIIs and for LCN1 reveal the presence of the corresponding mRNAs in the placenta. The techniques of immunohistochemical detection by revelation using peroxidase and DAB reveal the presence of the hOBPII (figure 19) and LCN1 (figure 20) proteins very strongly in the lacunae of the placental villi and much more weakly, and very diffusely, in the epithelial cells of the amniotic membrane and in subjacent cells, both for the amniotic membrane and for the cord. The expression of the mRNAs detected by RT-PCR in the placental villi indicates that these hOBPII proteins are produced by the placenta in order to transfer lipophilic molecules (hormones, fatty acids, vitamins, etc.) to the fetus via the fetal blood circulation. The proteins detected in the amniotic membrane may correspond to the absorption of these proteins from the amniotic fluid, since it is known that the nasal structures of the fetus are bathed in this amniotic fluid. The hOBPII and LCN1 proteins would thus contribute to the transport of lipophilic substances which are small in size from the mother to the fetus in the chorionic villi via the fetal blood, and from the fetus to the mother via the amniotic fluid.

The RT-PCRs carried out using the sets of primers for the ApoD protein reveal the presence of the corresponding mRNAs in the placenta. The techniques of immunohistochemical detection by revelation using peroxidase and DAB reveal the presence of the ApoD protein very strongly in the lacunae of the placental villi (figure 21). The ApoD protein is not detected in the cells of the amniotic membrane nor in subjacent cells, both for the amniotic membrane and for the cordon. The expression of the mRNAs detected by RT-PCR

in the placenta indicates that the ApoD protein is produced by the placenta in order to transfer lipophilic molecules (hormones, fatty acids, vitamins, etc.) to the fetus via the fetal blood circulation. The ApoD protein may also be produced by the fetal liver. The ApoD protein would thus contribute to the transport of lipophilic substances which are small in size from the mother to the fetus in the chorionic villi via the fetal blood.

More generally, the expression of lipocalins by the placenta, such as alpha-1-acid glycoprotein (A1AG), prostaglandin D synthase (PGDS), PP14 and RBP, as revealed by the RT-PCR techniques, indicates that these proteins, which have hydrophobic molecule-transporting properties, are recruited in massive amounts for mother-child connections, for transporting these molecules which may be steroid hormones, fatty acids, vitamins or xenobiotics, and more generally metabolism-derived detoxification products.

REFERENCES

Attwood, J. *et al.* (1994) Genomics *19*, 203-214.

Barany, F., (1991), Proc. Natl. Acad. Sci. USA, 88, 189-193.

Beste *et al.* (1999), Proc. Natl. Acad. Sci. USA, 96, 1898-1903.

Bocskei, Z. *et al.* (1992) Nature *360*, 186-188.

Borghoff S.J. *et al.* (1990) Annu. Rev. pharmacol. Toxicol. 30, 349-367

Burg, J.L. et al. (1996), Mol. and Cell. Probes, 10, 257-271.

Chu, B.C.F. et al. (1986), Nucleic Acids Res., 14, 5591-5603.

Dear, T.N. *et al.* (1991) Biochemistry *30*, 10376-10382.

Dewald, G. *et al.* (1996) Ann Hum Genet *60*, 281-291.

Duck, P. *et al.* (1990), Biotechniques, *9*, 142-147.

Edwards, C.P., and Aruffo, A., (1993), Curr. Op. Biotechnology *4*, 558-563.

Erlich, H.A., (1989), New York: Stockton Press.

Flower, D.R. (1995) J Mol Recognit *8*, 185-195.

Flower, D.R. (1996) Biochem J *318*, 1-14.

Gilbertson T.A. (1998) Cur. Op. Neurobiology *8*, 447-452.

Gregoire, C. *et al.* (1996) J Biol Chem *271*, 32951-32959.

Guatelli J.C. *et al.* (1990), Proc. Natl. Acad. Sci. USA, 87, 1874-1878.

Henzel, W.J. *et al.* (1988) J Biol Chem *263*, 16682-16687.

Holzfeind, P. and Redl, B. (1994) Gene *139*, 177-183.

Hornigold, N. *et al.* (1997) Genomics *41*, 385-389.

Innis, M.A. et al. (1990), Academic Press.

Igarashi, M. *et al.* (1992) Proc Natl Acad Sci USA *89*, 5376-2380.

Kievitis, T. et al. (1991), J. Virol. Methods, 35, 273-286.

Kock, K. *et al.* (1994) Eur J Biochem *221*, 905-916.

Kohler, G. et al. (1975), Nature, 256 (5517), 495-497.

Konieczny, A. *et al.* (1997) Immunology *92*, 577-586.

Kwoh, D.Y. et al. (1989), Proc. Natl. Acad. Sci. USA, 86, 1173-1177.

Lacazette, E. *et al.* (1997) Ann Hum Genet *61*, 449-455.

Landegren, U. *et al.* (1988), Science, 241, 1077-1080.

Lassagne, H. and Gachon, A.M. (1993) Exp Eye Res *56*, 605-609.

Lassagne, H. *et al.* (1993) Genomics *18*, 160-161.

Lassagne, H. *et al.* (1995) Cytogenet Cell Genet *71*, 104.

Lizardi, P.M. *et al.* (1988) Biotechnology *6*, 1197-1202.

Luckow, V.A. (1993) Curr. Op. Biotechnology 4, 564-572.

Matthews, J.A. *et al.* (1988), Anal. Biochem. 169: 1-25.

Miele, E.A. *et al.* (1983) J. Mol. Biol. 171: 281-295.

Miller, W.L. (1998) Clin Perinatol *25*, 799-817, v.

Monaco, H.L. and Zanotti, G. (1992) Biopolymers *32*, 457-465.

Nagata, A. *et al.* (1991) Proc Natl Acad Sci USA *88*, 4020-4024.

Nahmias, J. *et al.* (1995) Eur J Hum Genet *3*, 65-77.

Pelosi, P. (1996) J Neurobiol *30*, 3-19.

Pervaiz, S. and Brew, K. (1987) Faseb J *1*, 209-214.

Redl, B. *et al.* (1992) J Biol Chem *267*, 20282-20287.

Rolfs, A. *et al.* (1991), Berlin: Springer-Verlag.

Sambrook, J. *et al.* (1989), T. Sec. Ed. Cold Spring Harbor Lab., Cold Spring Harbor, New York.

Segev, D., (1992), Kessler C. Springer Verlag, Berlin, New York, 197-205.

Senoo, H. *et al.* (1990) J Lipid Res *31*, 1229-1239.

Shaw, P.H. *et al.* (1983) Cell *32*, 755-761.

Simard *et al.* (1992) Endocrinology *30*, 1115-1121.

Spinelli, S. *et al.* (1998) Biochemistry *37*, 7913-7918.

Stoesz S.P. and Gould M.N. (1995) Oncogene 11: 2233-2241.

Stone, B.B. *et al.* (1996) Mol. and Cell. Probes 10: 359-370.

van Slegtenhorst, M. *et al.* (1995) Eur J Hum Genet *3*, 78-86.

Walker, G.T. *et al.* (1992) Nucleic Acids Res. 20: 1691-1696.

Zeng, C. *et al.* (1996) Proc Natl Acad Sci USA *93*, 6626-6630.

CLAIMS

1. An isolated polypeptide comprising an amino acid sequence having at least 90% identity with the amino acid sequences SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 12 or SEQ ID No. 14.

2. An isolated polypeptide, characterized in that it comprises a polypeptide chosen from:

a) a polypeptide of sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 12 or SEQ ID No. 14;

b) a polypeptide which is a variant of a polypeptide of amino acid sequences defined in a);

c) a polypeptide homologous to the polypeptide defined in a) or b) and comprising at least 90% identity with said polypeptide of a);

d) a fragment of at least 15 consecutive amino acids of a polypeptide defined in a), b) or c);

e) a biologically active fragment of a polypeptide defined in a), b) or c).

3. An isolated polypeptide selected from a polypeptide corresponding to the sequence SEQ ID No. 2 and named $OBPII_{a\alpha}$, to the sequence SEQ ID No. 4 and named $OBPII_{a\beta}$, to the sequence SEQ ID No. 6 and named $OBPII_{a\gamma}$, and to the sequence SEQ ID No. 12 and named $OBPII_{b\beta}$.

4. The polypeptide as claimed in any one of claims 1 to 3, characterized in that it comprises at least the Gly-Thr-Trp-Tyr domain.

5. An isolated polynucleotide, characterized in that it encodes a polypeptide as claimed in any one of claims 1 to 4.

6. The polynucleotide as claimed in claim 5, of sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 11 or SEQ ID No. 13.

7. An isolated polynucleotide, characterized in that it comprises a polynucleotide chosen from:

a) a polynucleotide of sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 11 or SEQ ID No. 13, or the sequence of which is that of the RNA corresponding to the sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 11 or SEQ ID No. 13;

b) a polynucleotide, the sequence of which is complementary to the sequence of a polynucleotide defined in a),

c) a polynucleotide, the sequence of which comprises at least 80% identity with a polynucleotide defined in a) or b),

d) a fragment of at least 15 consecutive nucleotides, in preference 21 consecutive nucleotides, and preferably 30 consecutive nucleotides, of a polynucleotide defined in a), b), c) or d).

8. The polynucleotide as claimed in claim 7, characterized in that it is labeled directly or indirectly with a radioactive compound or a nonradioactive compound.

9. The use of a polynucleotide as claimed in claim 8, as a probe for detecting nucleic acid sequences.

10. The use of a polynucleotide as claimed in any one of claims 5 to 8, as a primer for amplifying or polymerizing nucleic acid sequences.

11. The use of a polynucleotide as claimed in claim 8, as a sense or antisense oligonucleotide for controlling the expression of the corresponding protein product.

12. A recombinant vector for cloning a polynucleotide as claimed in one of claims 5 to 8 and/or for expressing a polypeptide as claimed in one of claims 1 to 4, characterized in that it contains a polynucleotide as claimed in any one of claims 5 to 8.

13. The vector as claimed in claim 12, characterized in that it comprises the elements which allow the expression, optionally the secretion, of said polypeptide in a host cell.

14. The vector as claimed in any one of claims 12 to 13, characterized in that the elements which allow the expression of said polypeptide are chosen from:

a) the isolated polynucleotide of sequence SEQ ID No. 15 and SEQ ID No. 16;

b) a polynucleotide, the sequence of which is complementary to the sequence of the polynucleotide defined in a);

c) a polynucleotide, the sequence of which comprises at least 80% identity with a polynucleotide defined in a) or in b);

d) a polynucleotide which hybridizes, under conditions of high stringency, with a sequence of the polynucleotide defined in a), b) or c).

15. The vector as claimed in claims 13 and 14, for expression in eukaryotic cells, selected from viral DNA and naked DNA.

16. A host cell, characterized in that it is transformed with a vector as claimed in one of claims 12 to 15.

17. A process for preparing a recombinant polypeptide, characterized in that a host cell as claimed in claim 16 is cultured under conditions which allow the expression and, optionally, the secretion of said recombinant polypeptide, and in that said recombinant polypeptide is recovered.

18. A recombinant polypeptide which can be obtained using a process as claimed in claim 17.

19. The use of a polypeptide chosen from a polypeptide as claimed in any one of claims 1 to 4 and 18, or a fragment thereof, as a protein which binds to a hydrophobic ligand, preferably an odorous molecule.

20. The use of a polypeptide as claimed in any one of claims 1 to 4 and 18, or of the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$, a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$ or a fragment thereof, as a competitive inhibitor, as an agonist or as an antagonist of the cellular receptors for lipocalins.

21. A monoclonal or polyclonal antibody, and fragments thereof, characterized in that it is specifically against an isolated polypeptide as claimed in one of claims 1 to 4 and 18.

22. The use of an antibody as claimed in claim 21, for demonstrating the presence of a polypeptide as claimed in one of claims 1 to 4 and 18 in a biological sample.

23. A process for detecting an antibody against hOBPII in human serum from an allergic and/or asthmatic patient, using an hOBPII polypeptide.

24. A process for controlling the volatilization of an odorant, characterized in that it comprises a step of binding of said odorant with a polypeptide as claimed in any one of claims 1 to 4 and 18, or with the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$.

25. The process as claimed in claim 24, characterized in that the polypeptide is bound to a solid support.

26. The process as claimed in claim 24, characterized in that the polypeptide is in a liquid composition.

27. The process as claimed in claim 26, characterized in that said composition is a fragranced composition for the skin.

28. A process for screening a molecule, preferably odorants or flavors, which comprises passing the molecule over a substrate which comprises a polypeptide as claimed in any one of claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, bound to said substrate, said polypeptide

binding said odorant or flavor, and recovering said odorant or flavor from the polypeptide if necessary.

29. The process as claimed in claim 28, characterized in that the molecules are human pheromones.

30. A process for solubilizing lipophilic molecules, characterized in that it comprises binding said molecule to a polypeptide as claimed in any one of claims 1 to 4 and 18, or to the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$.

31. The application of the polypeptides as claimed in one of claims 1 to 4 and 18, or of the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, of a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or of a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, in combination with dietary fatty acids, as a food supplement.

32. The application as claimed in claim 31, for preparing a medicinal product intended for the treatment of hyperlipidemias and obesity.

33. The application as claimed in claim 31, for supplementing nonmaternal milks.

34. The polypeptide as claimed in any one of claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, as an agent for targeting a pharmaceutical compound.

35. The polypeptide as claimed in claim 34, characterized in that said polypeptide is expressed in the form of a protein from fusion with a protein which allows specific cellular addressing.

36. The polypeptide as claimed in claim 35, characterized in that said protein which allows specific cellular addressing is chosen from the group composed of interleukins, of cytokines, of lymphokines, of interferons, of growth factors, of hormones and of antibodies.

37. The polypeptide as claimed in claim 34, characterized in that said polypeptide is combined with a molecule which allows specific cellular addressing.

38. The polypeptide as claimed in claim 37, characterized in that said molecule which allows specific cellular addressing is chosen from the group composed of steroids, of interleukins, of cytokines, of lymphokines, of interferons, of growth factors, of hormones and of antibodies.

39. The polypeptide as claimed in any one of claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, as a pharmaceutical compound transporter.

40. A pharmaceutical composition comprising a pharmaceutical compound bound at least to a polypeptide as claimed in one of claims 34 to 39, and a pharmaceutically acceptable vehicle.

41. The pharmaceutical composition as claimed in claim 40, characterized in that the pharmaceutical compound is chosen from the group of anticancer agents.

42. The pharmaceutical composition as claimed in claim 41, characterized in that said anticancer agent is a radioactive isotope chosen from the group: iodine^{131}, yttrium^{90}, gold^{199}, palladium^{100}, copper^{67}, bismuth^{217} and antimony^{211}.

43. The pharmaceutical composition as claimed in any one of claims 40 to 42, characterized in that said polypeptide as claimed in any one of claims 34 to 39 constitutes a delayed form of delivery of said pharmaceutical compound in the body.

44. A pharmaceutical composition comprising an expression vector as claimed in claim 12 or 13, and a pharmaceutically acceptable vehicle.

45. The pharmaceutical composition as claimed in any one of claims 40 to 44, for treating cancer preferably chosen from breast cancer, uterine cancer, prostate cancer, liver cancer and pulmonary epithelial cell carcinoma.

46. The use of a polypeptide chosen from the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ and a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, for preparing a medicinal product intended for the treatment of uterine cancer, prostate cancer, liver cancer and pulmonary epithelial cell cancer.

47. The use of the polypeptide as claimed in any one of claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of

sequence SEQ ID No. 10, or a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, for preparing a medicinal product intended for the transport of a compound across the placental barrier.

48. The use as claimed in claim 47, characterized in that said transport is carried out from the gestating mother to the fetus, and in that said compound is chosen from hormones, essential fatty acids, lipophilic medicinal products and vitamins.

49. The use as claimed in claim 47, characterized in that said transport is carried out from the fetus to the mother and is intended for the detoxification of the fetus.

50. The polypeptide as claimed in claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, as a pregnancy marker.

51. The polypeptide as claimed in claims 1 to 4 and 18, or the polypeptide $hOBPII_{b\alpha}$ of sequence SEQ ID No. 10, a polypeptide which is a variant of said polypeptide $hOBPII_{b\alpha}$ or a polypeptide homologous to said polypeptide $hOBPII_{b\alpha}$ comprising at least 90% identity with said polypeptide $hOBPII_{b\alpha}$, as a marker for a feto-placental pathological condition.

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

(43) Date de la publication internationale
22 février 2001 (22.02.2001)

PCT

(10) Numéro de publication internationale
WO 01/12806 A2

(51) Classification internationale des brevets⁷: C12N 15/12, C07K 14/47, C12Q 1/68, C12N 15/11, 5/10, C07K 16/18, G01N 33/68, A61K 7/32, 7/46, A23C 9/152, C07K 19/00 // A61P 3/04, 3/06, 35/00

(21) Numéro de la demande internationale:
PCT/FR00/02319

(22) Date de dépôt international: 11 août 2000 (11.08.2000)

(25) Langue de dépôt: français

(26) Langue de publication: français

(30) Données relatives à la priorité:
99/10439 12 août 1999 (12.08.1999) FR

(71) Déposant (*pour tous les États désignés sauf US*): UNIVERSITE D'AUVERGNE [FR/FR]; 49, boulevard François-Mitterrand, F-63000 Clermont-Ferrand (FR).

(71) Déposant et
(72) Inventeur: PITIOT, Gilles [FR/FR], 151, rue du Chevaleret, F-75013 Paris (FR).

(72) Inventeurs; et
(75) Inventeurs/Déposants (*pour US seulement*): LACAZETTE, Eric [FR/FR]; Appartement 211, 66, boulevard La Fayette, F-63000 Clermont-Ferrand (FR). GACHON, Françoise [FR/FR]; 11, rue des Paillards, F-63540 Romagnat (FR).

(74) Mandataires: MARTIN, Jean-Jacques etc.; Cabinet Regimbeau, 26, avenue Kléber, F-75116 Paris (FR).

(81) États désignés (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) États désignés (*régional*): brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Publiée:
— *Sans rapport de recherche internationale, sera republiée dès réception de ce rapport*

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT

(54) Title: ODORANT BINDING HUMAN PROTEINS FIXING HYDROPHOBIC LIGANDS: POLYPEPTIDES AND POLYNUCLEOTIDES CODING FOR SAID POLYPEPTIDES AND USES THEREOF

(54) Titre: "ODORANT-BINDING" PROTEINES HUMAINES FIXANT DES LIGANDS HYDROPHOBES: POLYPEPTIDES ET POLYNUCLEOTIDES CODANT LESDITS POLYPEPTIDES ET LEURS APPLICATIONS

(57) Abstract: The invention concerns novel polypeptides fixing hydrophobic ligands and in particular odours, called OPBII (Odorant Binding Protein), the polynucleotides coding for said polypeptides and specific antibodies directed against said polypeptides. The invention also concerns the uses of said molecules for personal hygiene, uses in agri-food system and nutritional and therapeutic uses

(57) Abrégé: L'invention concerne de nouveaux polypeptides fixant des ligands hydrophobes et notamment des odeurs, dénommés OBPII (Odorant Binding Protein), les polynucléotides codant pour lesdits polypeptides et des anticorps spécifiques dirigés contre lesdits polypeptides. L'invention concerne également les applications de ces molécules notamment pour l'hygiène corporelle, les applications agro-alimentaires, nutritionnelles et thérapeutiques.

WO 01/12806 A2

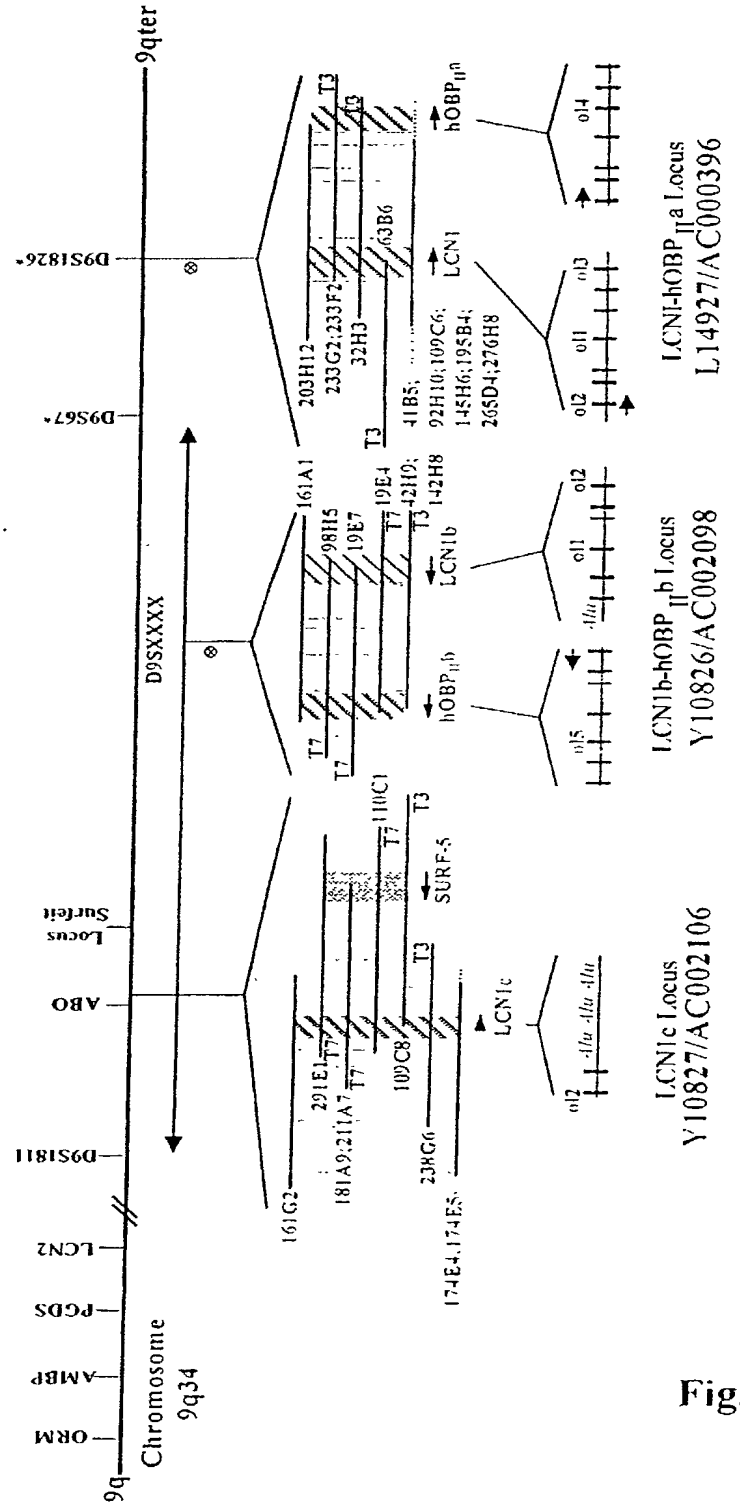

9qter
D9S1826
D9S67
D9SXXXX
Locus Surfeit
ABO
D9S1811
LCN2
PGDS
AMBP
ORM
9q
Chromosome 9q34
T3
T7
hOBP IIa
hOBP IIb
LCN1
LCN1b
LCN1c
SURF-5
203H12
233G2;233F2
32H3
63B6
41B5;
92H10;109C6;
145H6;195B4;
265D4;276H8
161A1
98H5
19E7
19E4
42H9;
142H8
111C1
291E1
181A9;211A7
100C8
238G6
161G2
174E4,174E5
LCN1-hOBP IIa Locus
L14927/AC000396
LCN1b-hOBP IIb Locus
Y10826/AC002098
LCN1c Locus
Y10827/AC002106

Fig. 1

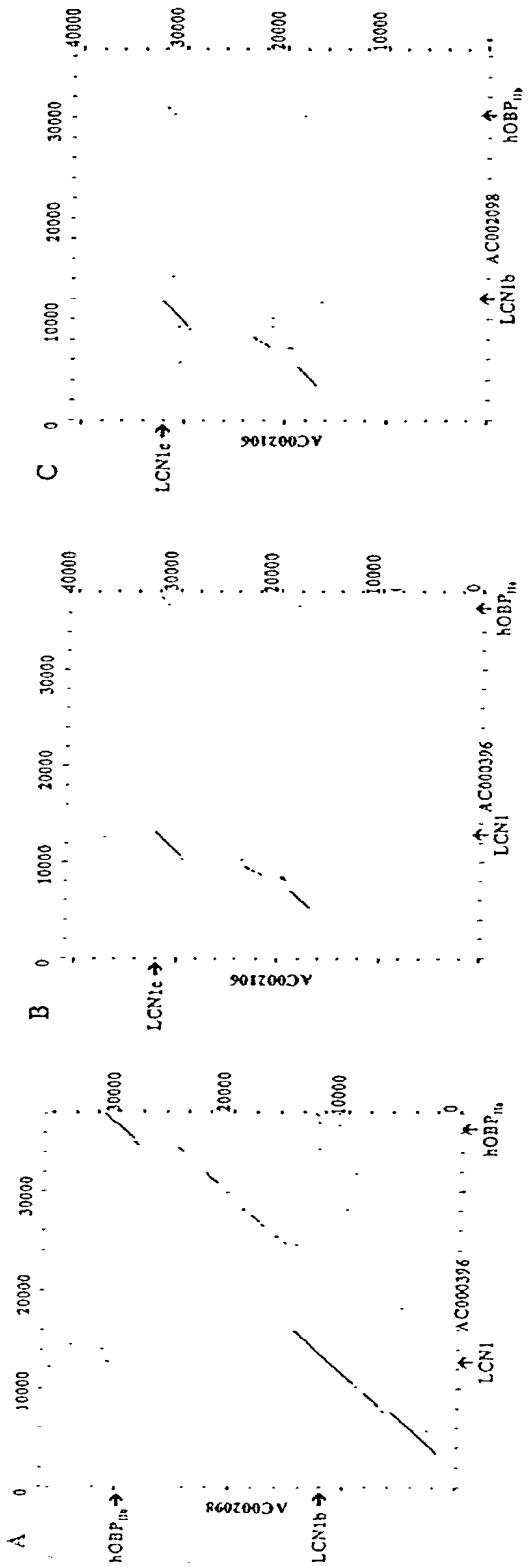

A
AC002098
AC000396
hOBP$_{IIa}$
LCN1b
LCN1
B
AC002106
AC000396
hOBP$_{IIa}$
LCN1c
LCN1
C
AC002106
AC002098
hOBP$_{IIa}$
LCN1c
LCN1b

FIG. 2

FIG. 3

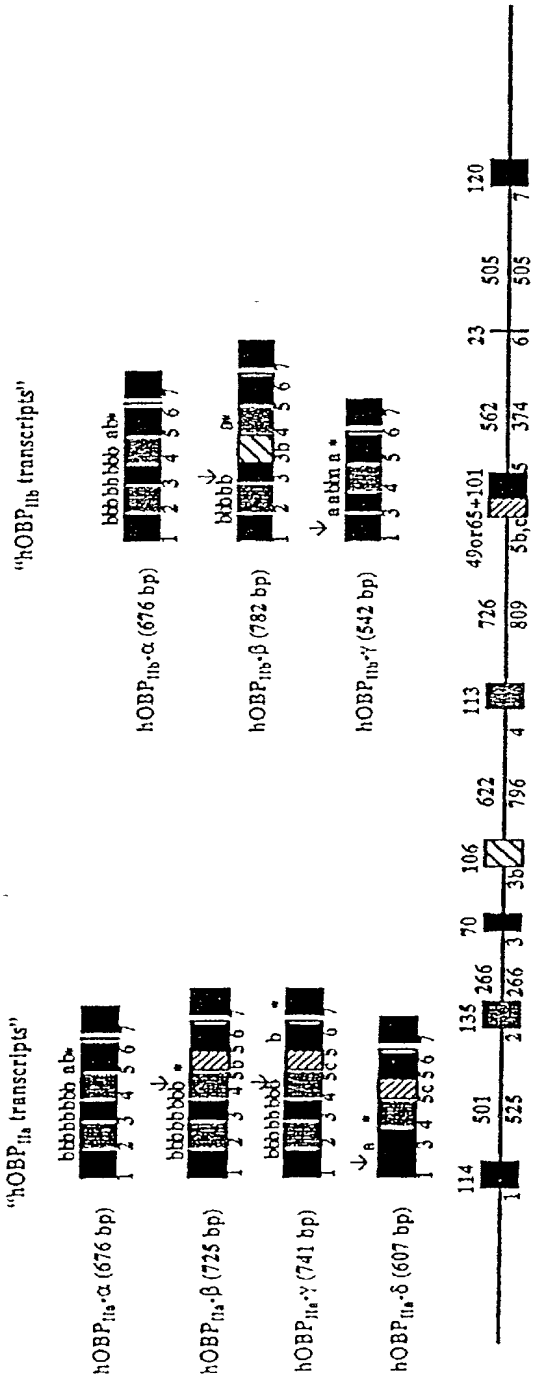
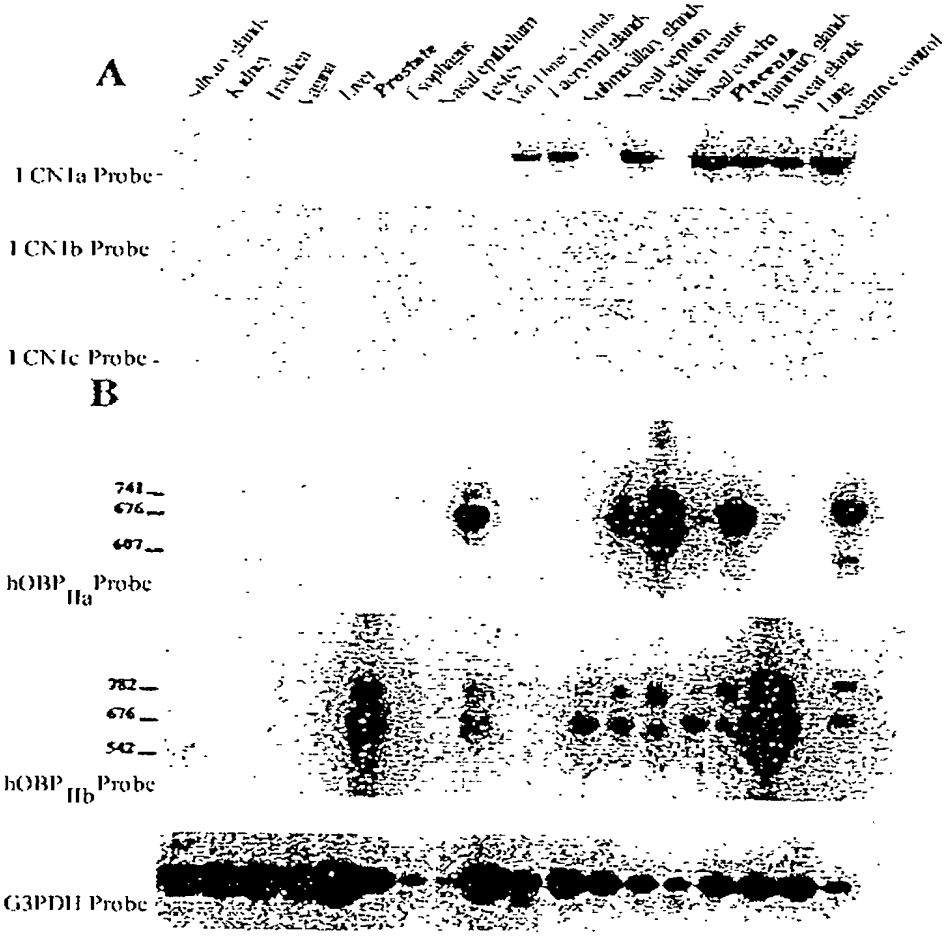

"hOBP11a transcripts"
"hOBP11b transcripts"
hOBP11a-α (676 bp)
hOBP11a-β (725 bp)
hOBP11a-γ (741 bp)
hOBP11a-δ (607 bp)
hOBP11b-α (676 bp)
hOBP11b-β (782 bp)
hOBP11b-γ (542 bp)
114
501
525
135
266
266
70
106
622
796
113
726
809
49or65+101
562
374
23
61
505
505
120

FIG.4

FIG.5

FIG.6


A
Kidney
Trachea
Liver
Prostate
Nasal epithelium
Testis
Submaxillary gland
Nasal septum
Middle meatus
Nasal cavity
Placenta
Mammary gland
Sweat glands
Lung
Negative control
hCNTa Probe
hCNTb Probe
hCNTc Probe
B
741
676
607
hOBP IIa Probe
782
676
542
hOBP IIb Probe
G3PDH Probe

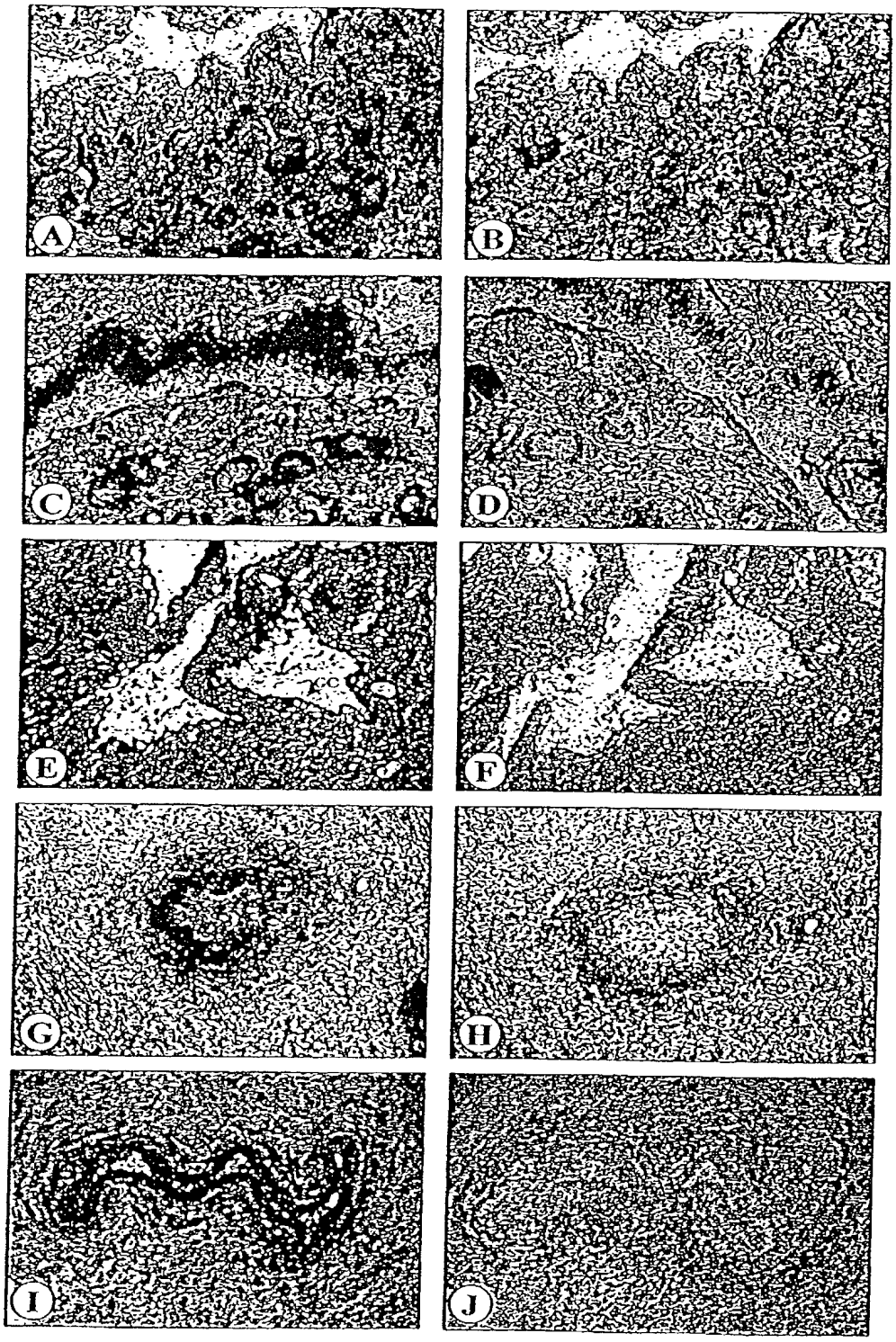


FIG. 7

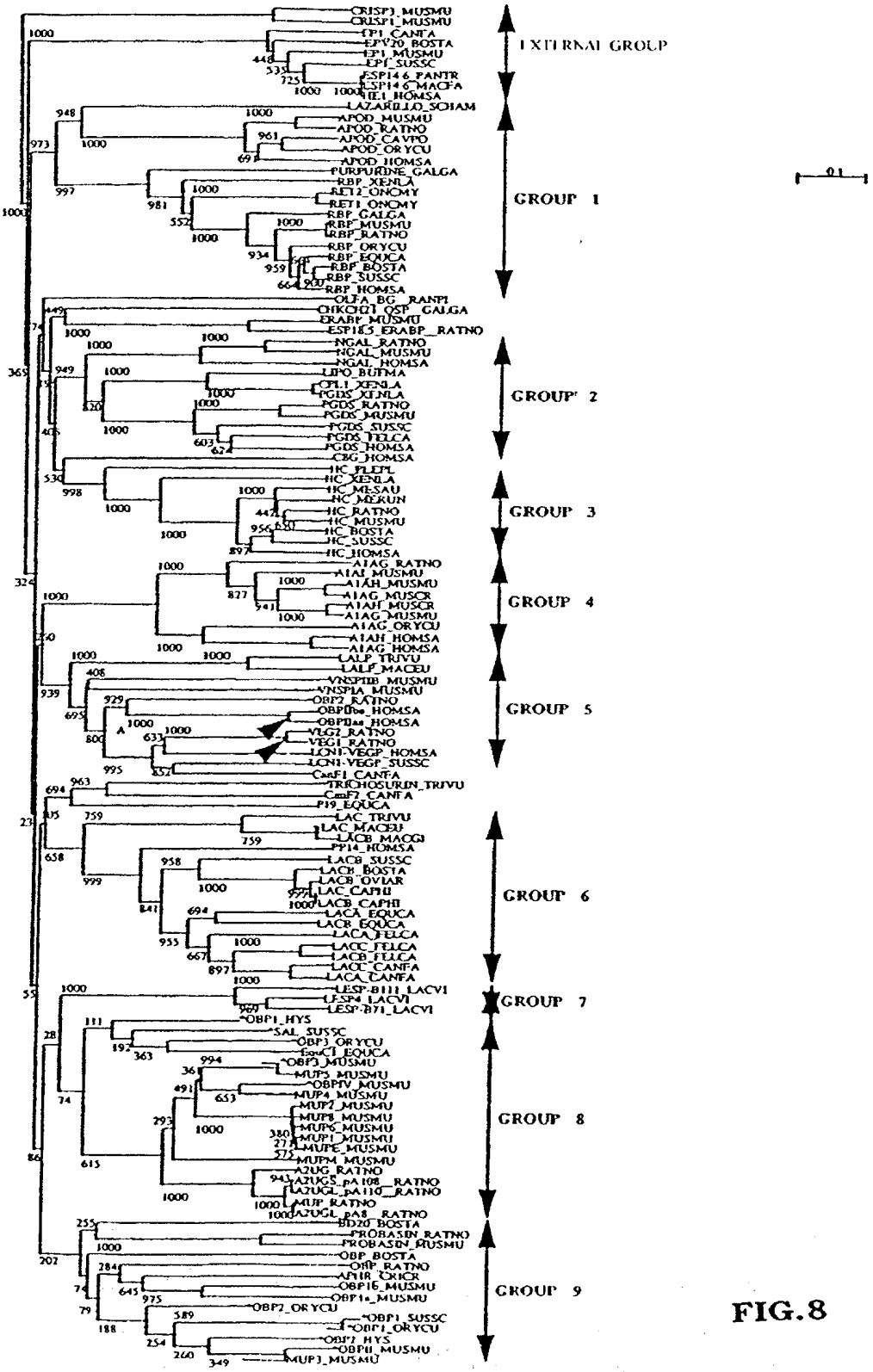

CRISP3_MUSMU
CRISP1_MUSMU
EP1_CANFA
EPV20_BOSTA
EP1_MUSMU
EPE_SUSSC
ESP146_PANTR
ESP146_MACFA
IE1_HOMSA
LAZARILLO_SCHAM
APOD_MUSMU
APOD_RATNO
APOD_CAVPO
APOD_ORYCU
APOD_HOMSA
PURPURINE_GALGA
RBP_XENLA
RET2_ONCMY
RET1_ONCMY
RBP_GALGA
RBP_MUSMU
RBP_RATNO
RBP_ORYCU
RBP_EQUCA
RBP_BOSTA
RBP_SUSSC
RBP_HOMSA
OLFA_BG_RANPI
OHKOGT_QSP_GALGA
ERABP_MUSMU
ESP185_ERABP_RATNO
NGAL_RATNO
NGAL_MUSMU
NGAL_HOMSA
LIPO_BUTMA
OPL1_XENLA
PGDS_XENLA
PGDS_RATNO
PGDS_MUSMU
PGDS_SUSSC
PGDS_FELCA
PGDS_HOMSA
CBG_HOMSA
HC_PLEPL
HC_XENLA
HC_MESAU
HC_MERUN
HC_RATNO
HC_MUSMU
HC_BOSTA
HC_SUSSC
HC_HOMSA
A1AG_RATNO
A1AI_MUSMU
A1AH_MUSMU
A1AG_MUSCR
A1AH_MUSCR
A1AG_MUSMU
A1AG_ORYCU
A1AH_HOMSA
A1AG_HOMSA
LALP_TRIVU
LALP_MACEU
VNSP1IB_MUSMU
VNSP1A_MUSMU
OBP2_RATNO
OBP1bo_HOMSA
OBP1aa_HOMSA
VEG2_RATNO
VEG1_RATNO
LCN1-VEGP_HOMSA
LCN1-VEGP_SUSSC
CmF1_CANFA
TRICHOSURIN_TRIVU
CmF7_CANFA
P19_EQUCA
LAC_TRIVU
LAC_MACEU
LACB_MAOGI
PP14_HOMSA
LACB_SUSSC
LACB_BOSTA
LACB_OVIAR
LAC_CAPHI
LACB_CAPHI
LACA_EQUCA
LACB_EQUCA
LACA_FELCA
LACC_FELCA
LACB_FELCA
LACC_CANFA
LACA_CANFA
LESP-B111_LACVI
LESP4_LACVI
LESP-B71_LACVI
OBP1_HYS
SAL_SUSSC
OBP3_ORYCU
6mC1_EQUCA
OBP3_MUSMU
MUP3_MUSMU
OBPIV_MUSMU
MUP4_MUSMU
MUP2_MUSMU
MUP8_MUSMU
MUP6_MUSMU
MUP1_MUSMU
MUPE_MUSMU
MUPM_MUSMU
A2UG_RATNO
A2UG5_pA108__RATNO
A2UGL_pA110__RATNO
MUP_RATNO
A2UGL_pA8__RATNO
BD20_BOSTA
PROBASIN_RATNO
PROBASIN_MUSMU
OBP_BOSTA
OBP_RATNO
AP1R_ORICR
OBP16_MUSMU
OBP1a_MUSMU
OBP2_ORYCU
OBP1_SUSSC
OBP1_ORYCU
OBP2_HYS
OBPII_MUSMU
MUP3_MUSMU
EXTERNAL GROUP
GROUP 1
GROUP 2
GROUP 3
GROUP 4
GROUP 5
GROUP 6
GROUP 7
GROUP 8
GROUP 9
0.1

FIG.8

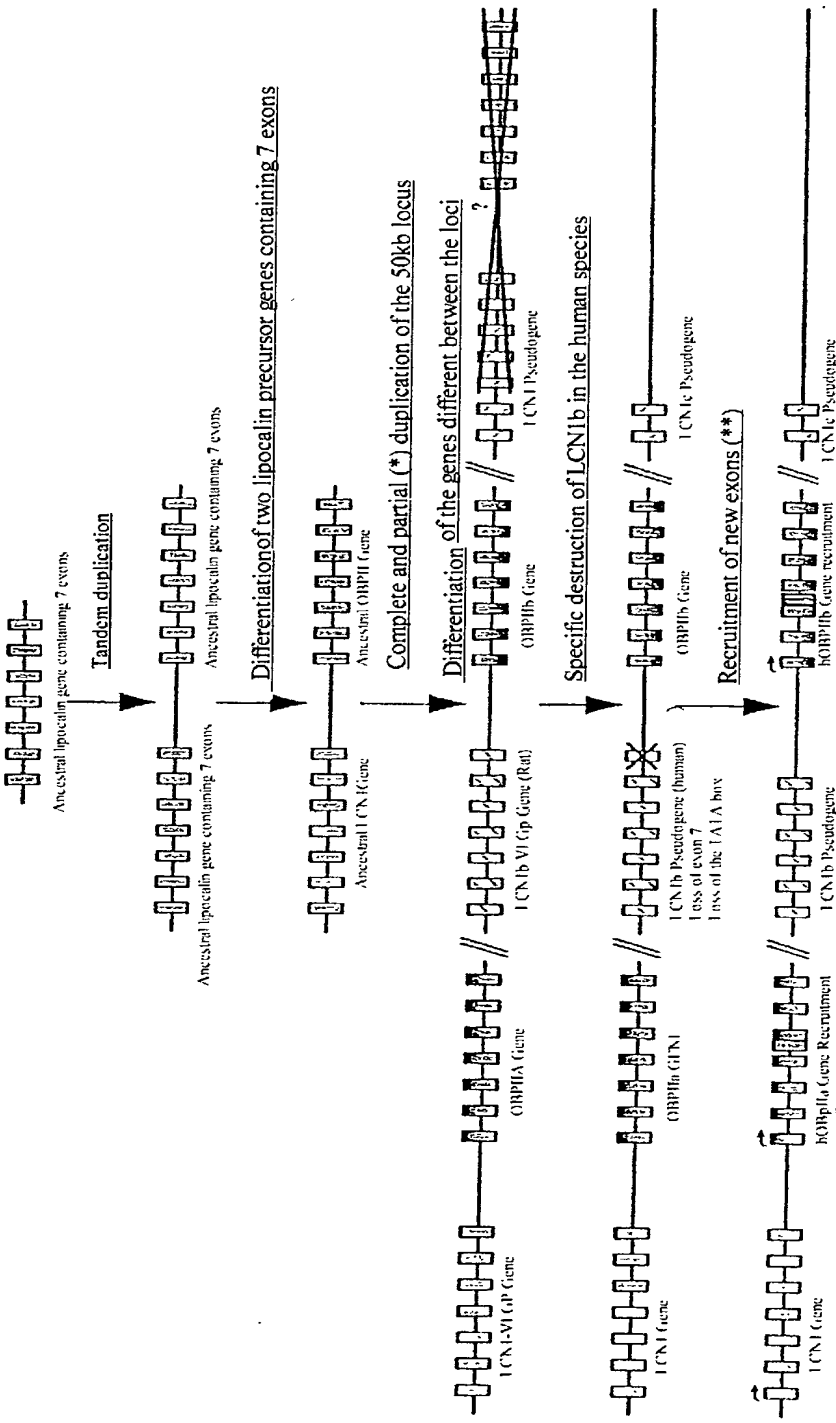

Ancestral lipocalin gene containing 7 exons
Tandem duplication
Ancestral lipocalin gene containing 7 exons
Ancestral lipocalin gene containing 7 exons
Differentiation of two lipocalin precursor genes containing 7 exons
Ancestral OBPII Gene
Ancestral LCN Gene
Complete and partial (*) duplication of the 50kb locus
Differentiation of the genes different between the loci
OBPIIb Gene
LCN Pseudogene
LCN1b VI Gp Gene (Rat)
OBPIIA Gene
LCN1c VI Gp Gene
Specific destruction of LCN1b in the human species
OBPIIb Gene
LCN1c Pseudogene
LCN1b Pseudogene (human)
Loss of exon 7
Loss of the TATA box
OBPIIa GENE
LCN1 Gene
Recruitment of new exons (**)
hOBPIIb Gene recruitment of an exon 3b
LCN1c Pseudogene
hOBPIIa Gene Recruitment of exons 4b and 5
LCN1b Pseudogene (human)
LCN1 Gene

FIG.9

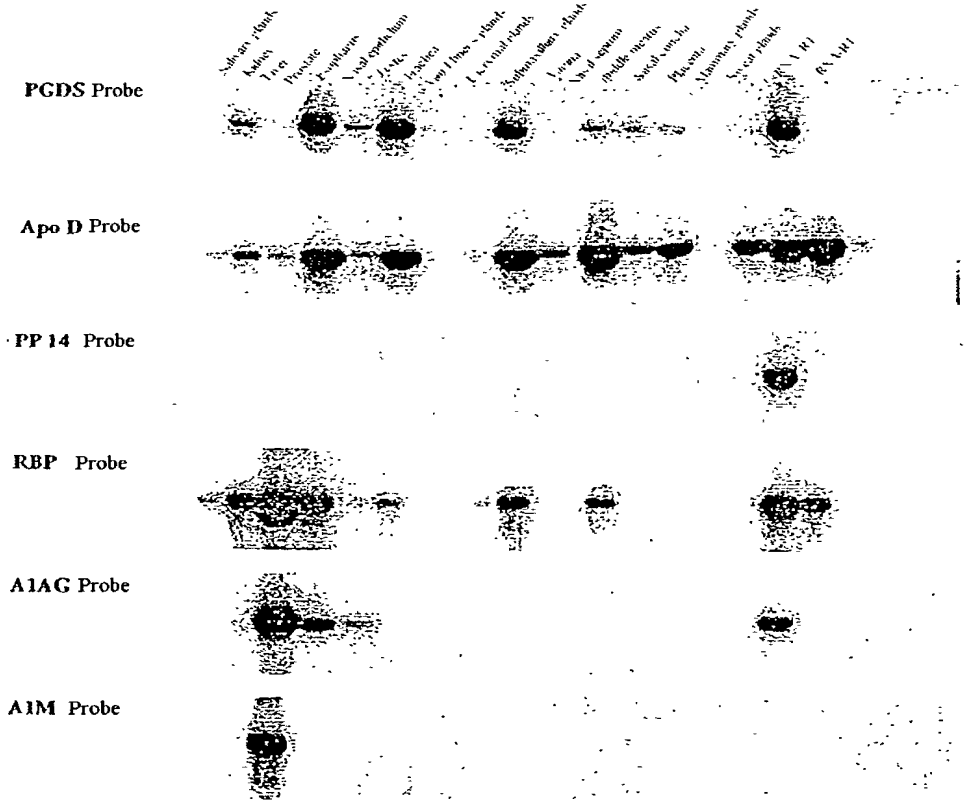

PGDS Probe
Apo D Probe
PP 14 Probe
RBP Probe
A1AG Probe
A1M Probe

FIG. 10

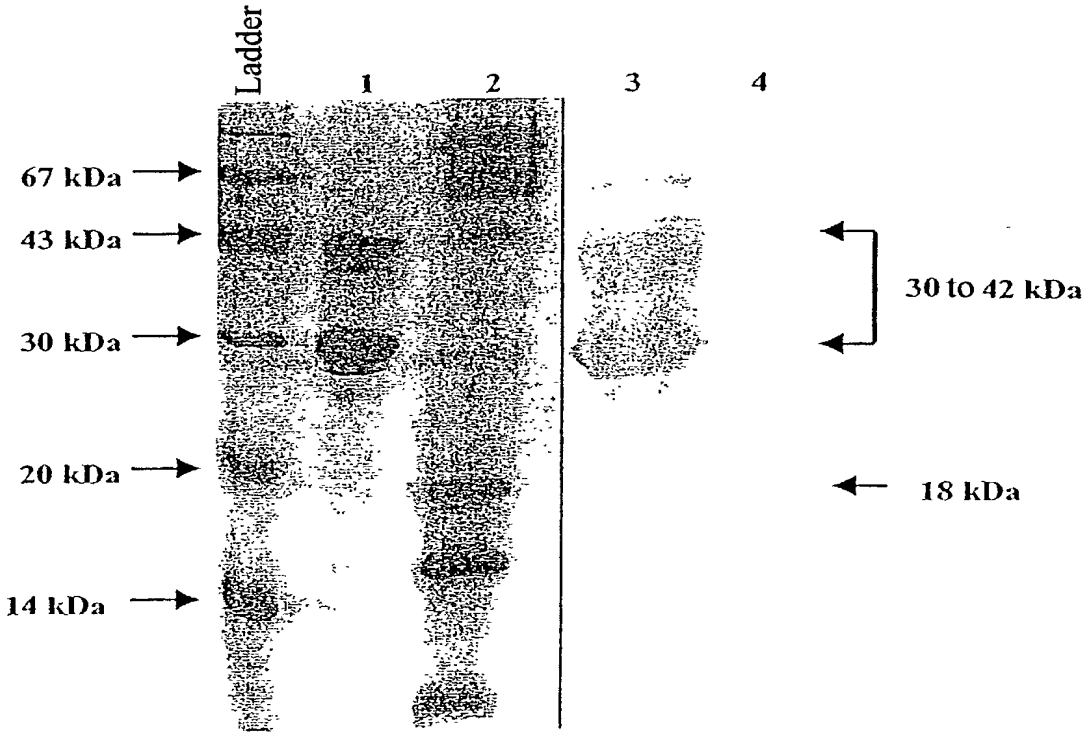


FIG.11

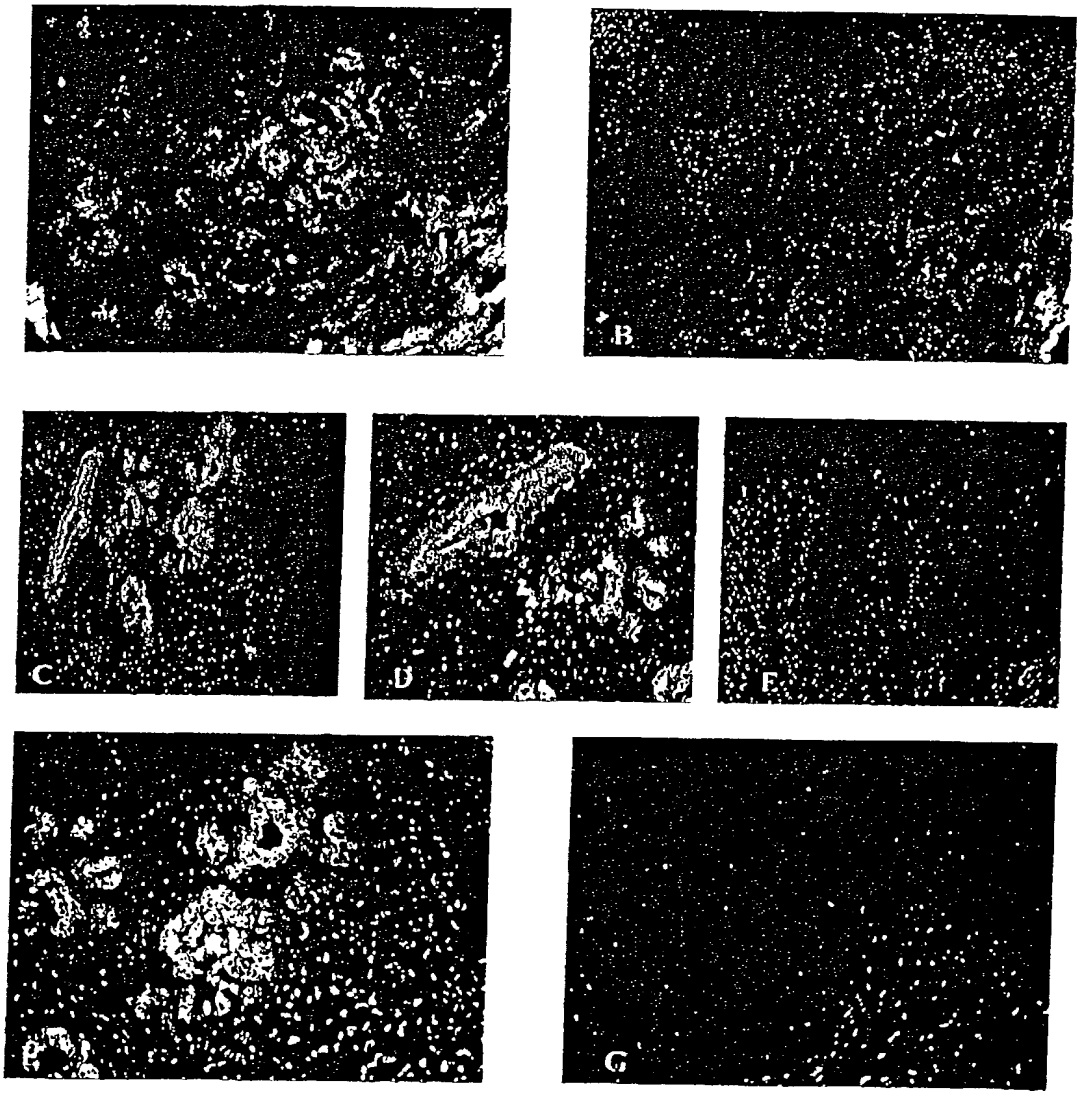


FIGURE 12

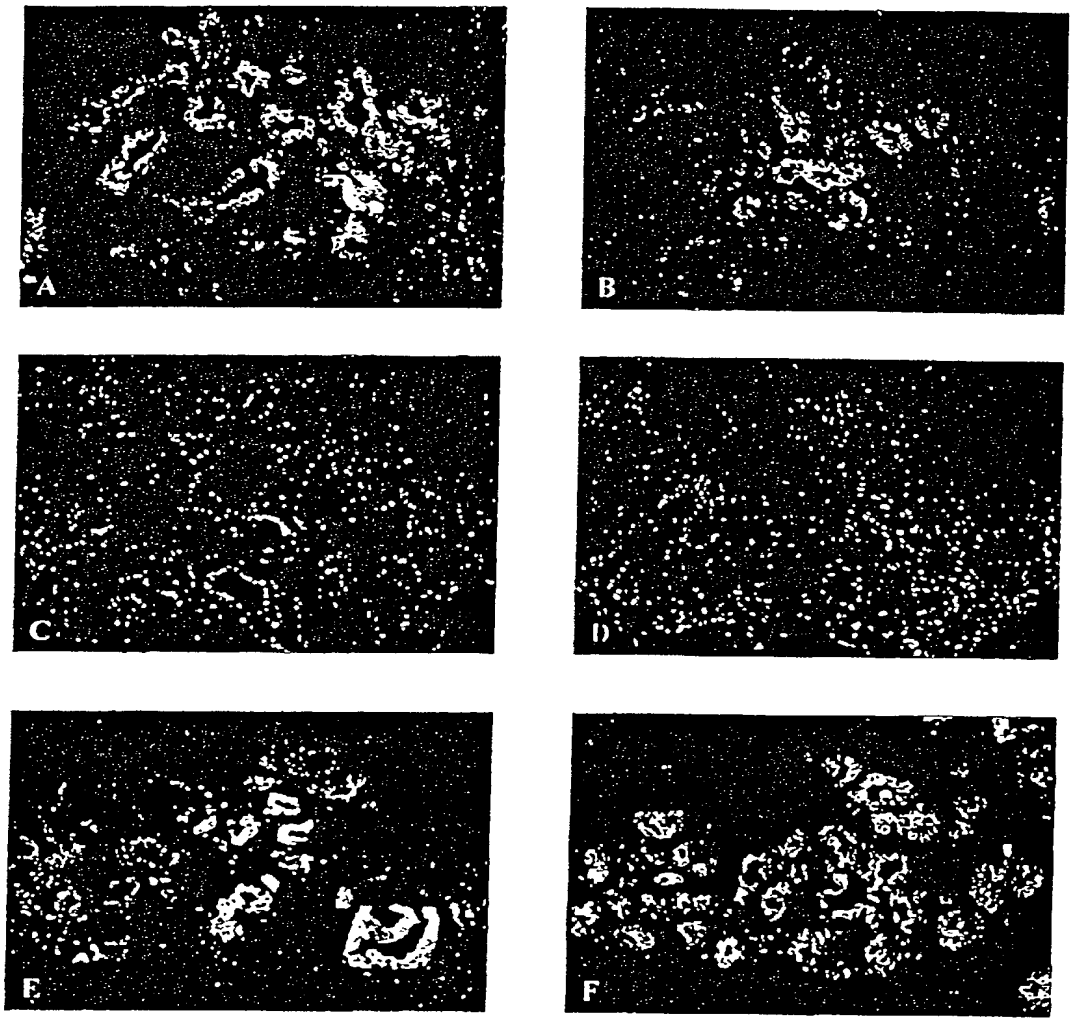


FIGURE 13

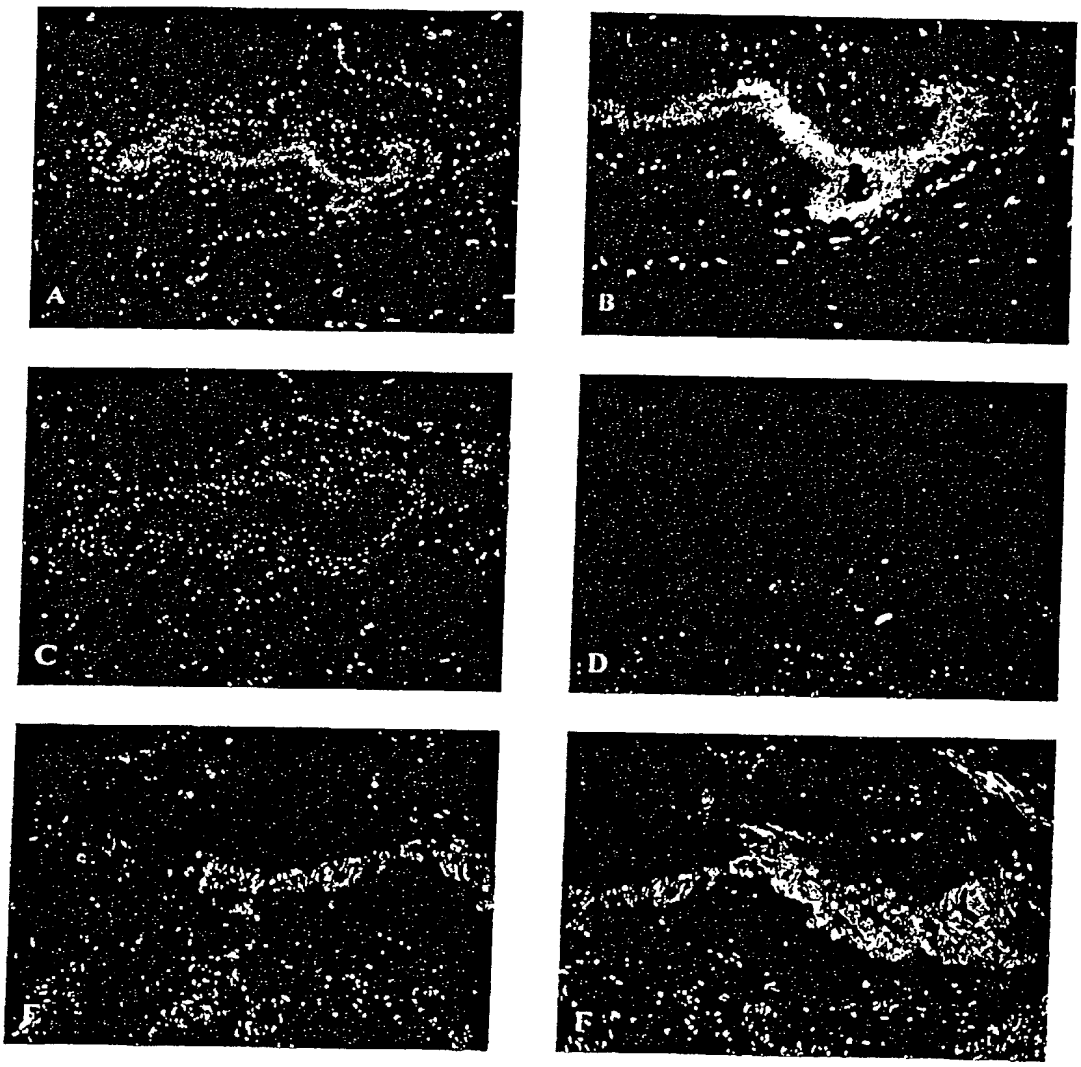


FIGURE 14

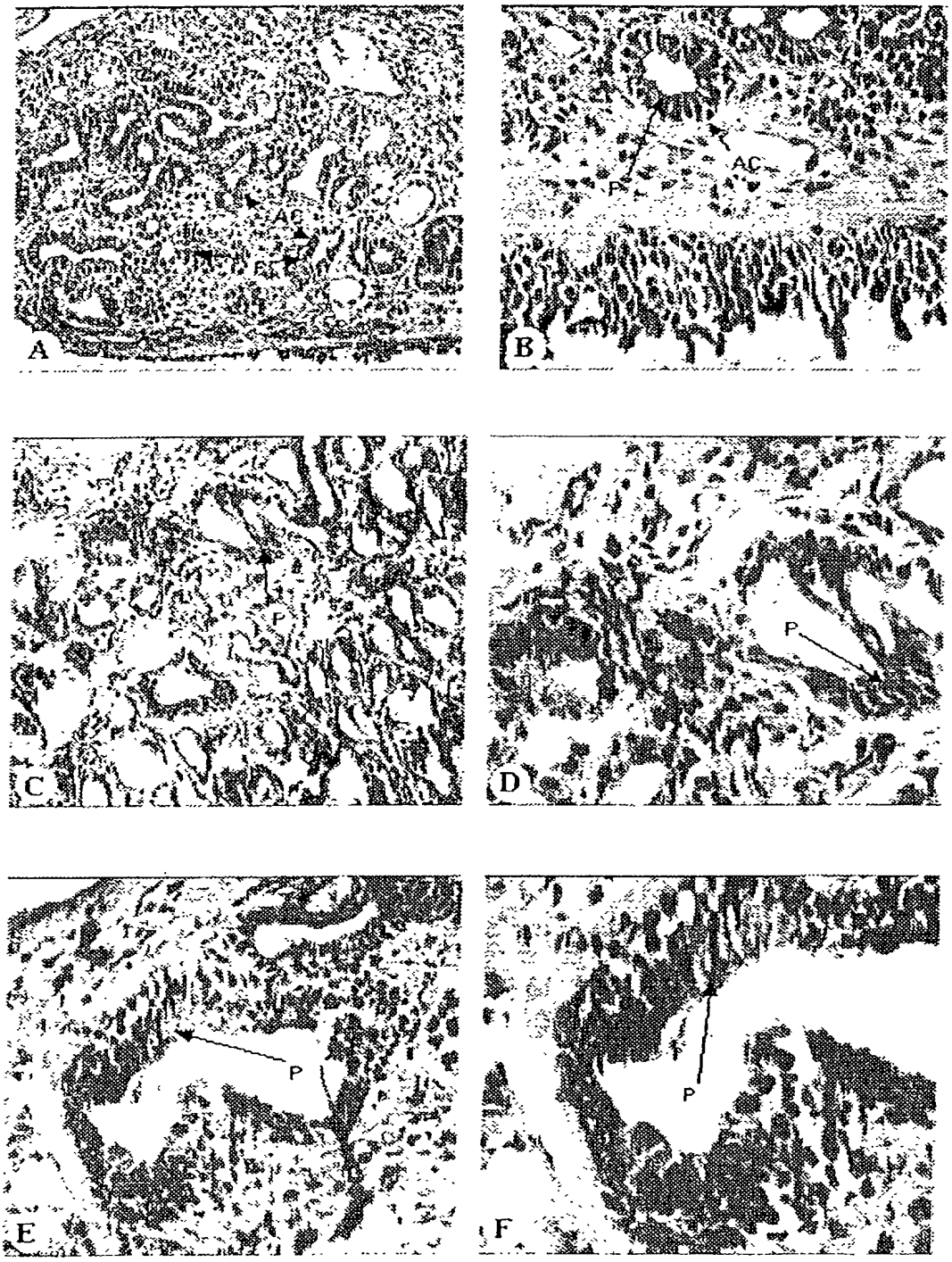


FIG. 15

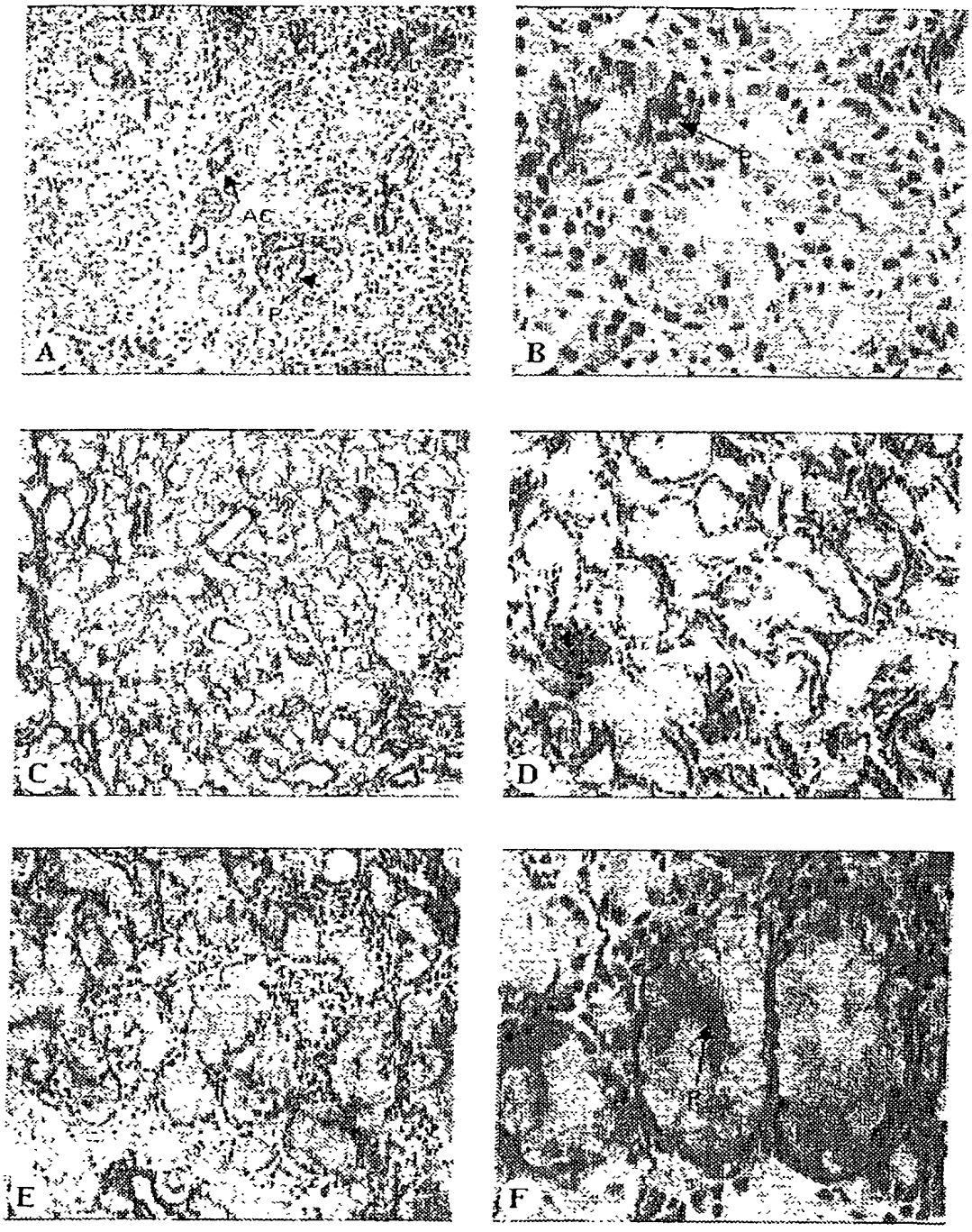


FIG. 16

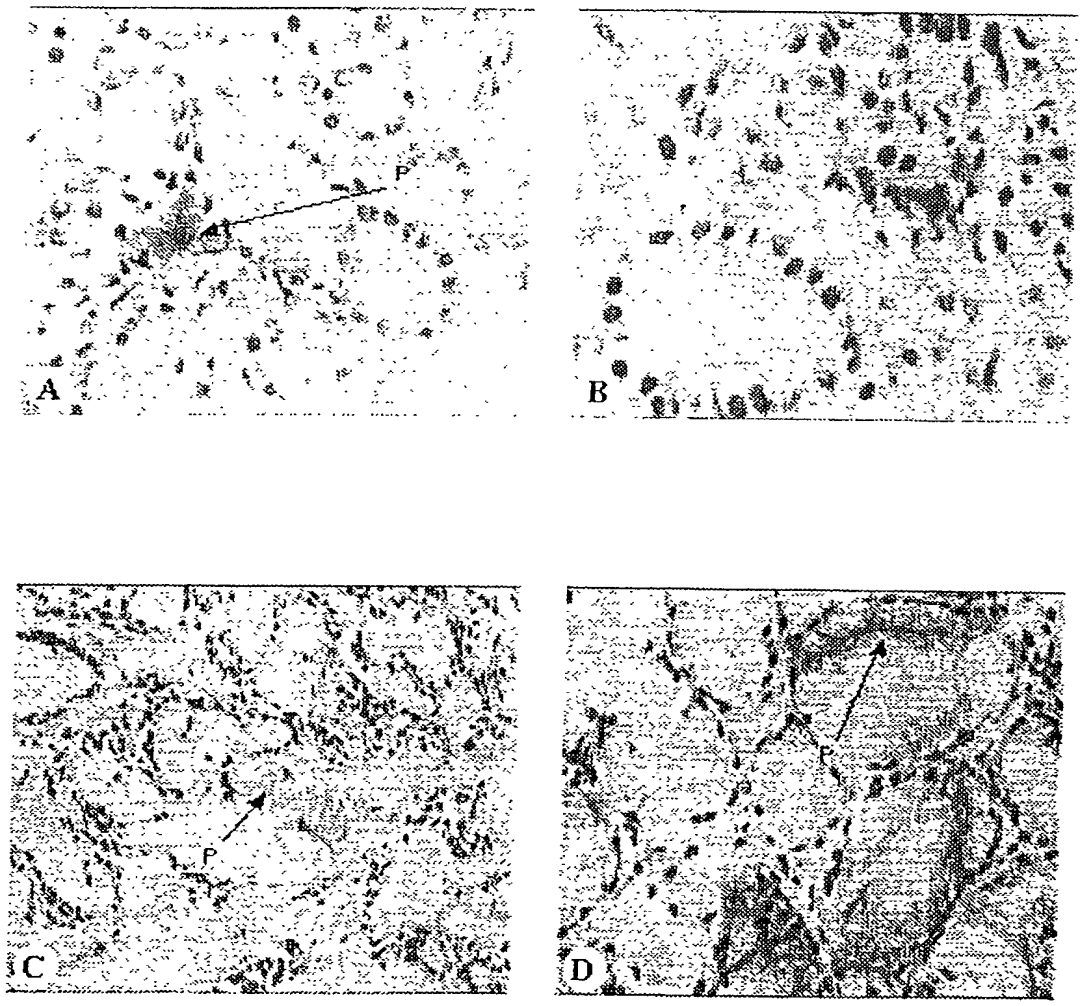


FIG. 17

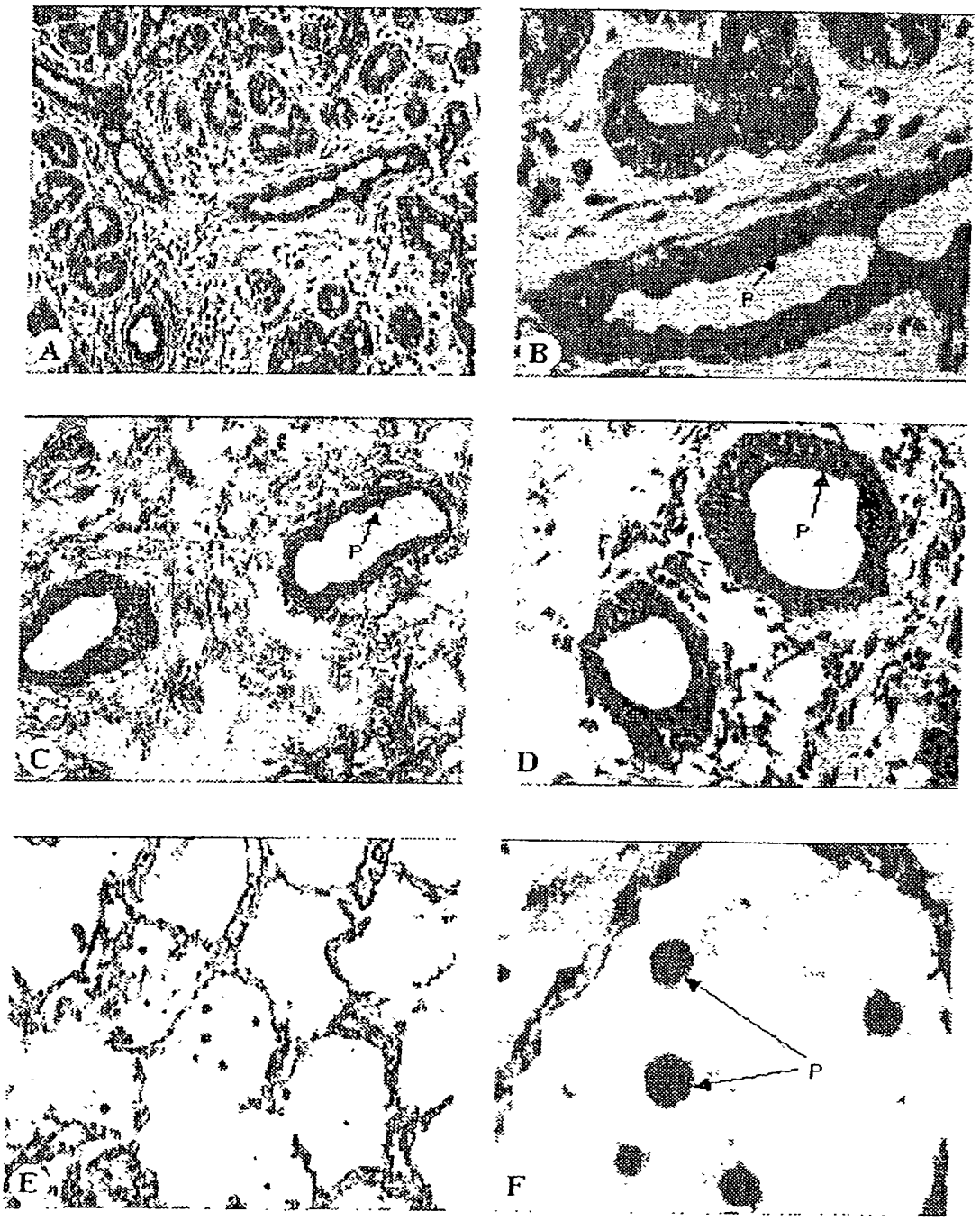


FIG. 18

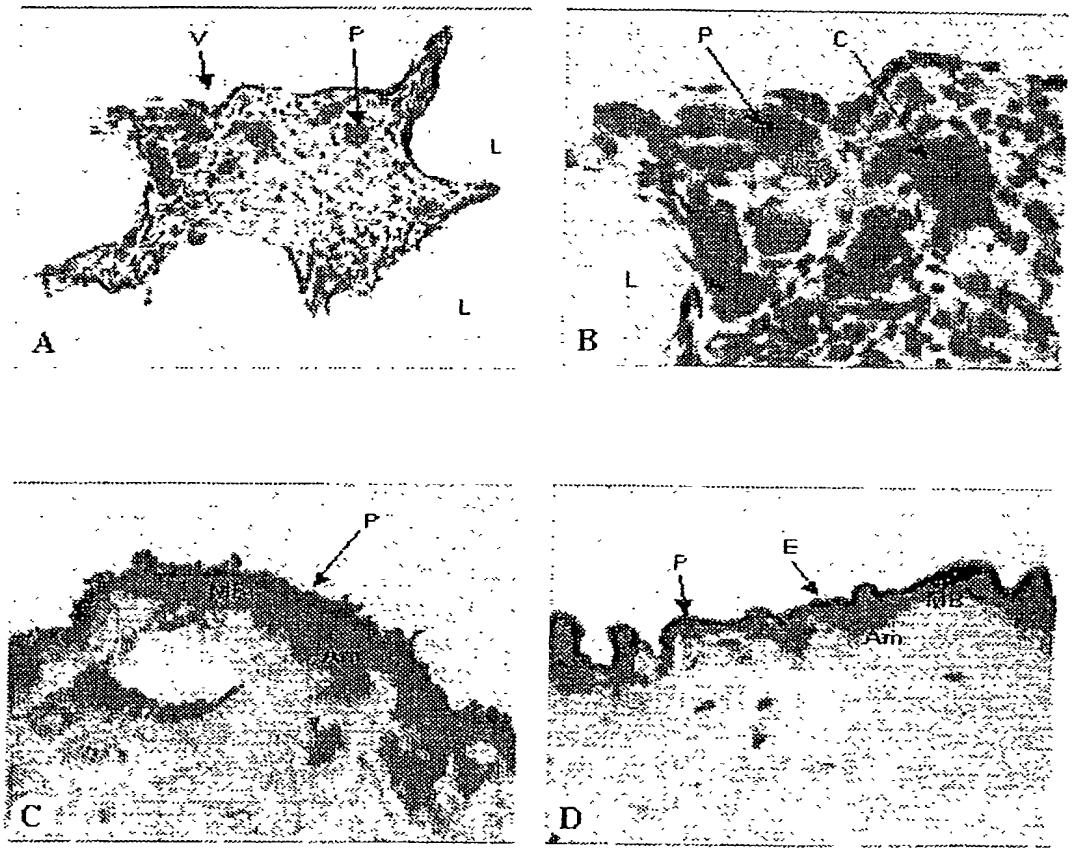


FIG. 19

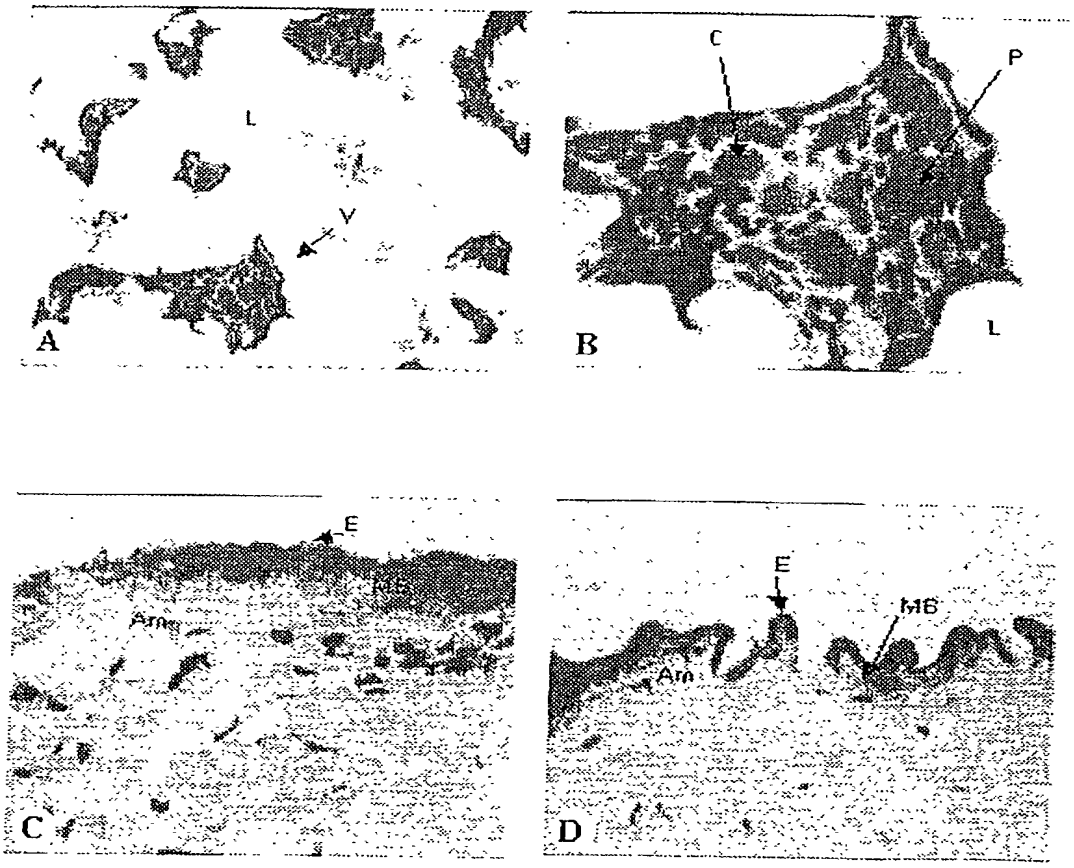


FIG. 20

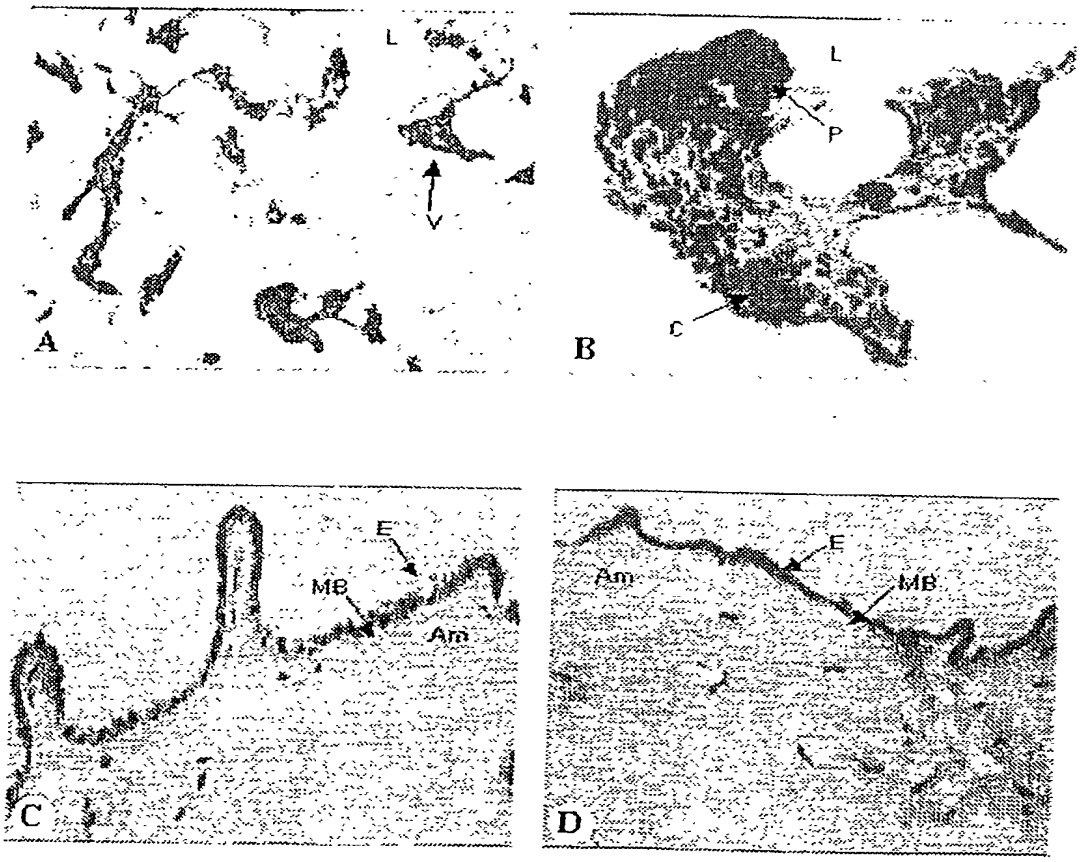


FIG. 21

Atty. Dkt. No. 065691-0260

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I HEREBY DECLARE:

THAT my residence, post office address, and citizenship are as stated below next to my name;

THAT I believe I am the original, first, and sole inventor (if only one inventor is named below) or an original, first, and joint inventor (if plural inventors are named below or in an attached Declaration) of the subject matter which is claimed and for which a patent is sought on the invention entitled **ODORANT-BINDING HUMAN PROTEINS HYDROPHOBIC LIGANDS: POLYPEPTIDES AND POLYNUCLEOTIDES CODING FOR SAID POLYPEPTIDES AND USES THEREOF**

(Attorney Docket No. 065691-0260)

the specification of which (check one)

____ is attached hereto.

XX was filed on Aug. 11, 2000 as ~~United States Application Number or~~ PCT International Application Number PCT/FR00/02319 and was amended on _____ (if applicable).

THAT I do not know and do not believe that the same invention was ever known or used by others in the United States of America, or was patented or described in any printed publication in any country, before I (we) invented it;

THAT I do not know and do not believe that the same invention was patented or described in any printed publication in any country, or in public use or on sale in the United States of America, for more than one year prior to the filing date of this United States application;

THAT I do not know and do not believe that the same invention was first patented or made the subject of an inventor's certificate that issued in any country foreign to the United States of America before the filing date of this United States application if the foreign application was filed by me (us), or by my (our) legal representatives or assigns, more than twelve months (six months for design patents) prior to the filing date of this United States application;

THAT I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment specifically referred to above;

THAT I believe that the above-identified specification contains a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the invention, and sets forth the best mode contemplated by me of carrying out the invention; and

THAT I acknowledge the duty to disclose to the U.S. Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I HEREBY CLAIM foreign priority benefits under Title 35, United States Code §119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below any foreign application for patent or inventor's certificate or of any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number	Country	Foreign Filing Date	Priority Claimed?	Certified Copy Attached?
99 10439	France	Aug. 12, 1999	yes	

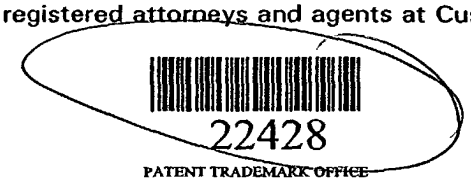
I HEREBY CLAIM the benefit under Title 35, United States Code § 119(e) of any United States provisional application(s) listed below.

U.S. Provisional Application Number	Filing Date

I HEREBY CLAIM the benefit under Title 35, United States Code, §120 of any United States application(s), or § 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application Number	PCT Parent Application Number	Parent Filing Date	Parent Patent Number
	PCT/FR00/02319	August 11, 2000	

I HEREBY APPOINT the registered attorneys and agents at Customer Number 22428



to have full power to prosecute this application and any continuations, divisions, reissues, and reexaminations thereof, to receive the patent, and to transact all business in the United States Patent and Trademark Office connected therewith.

I request that all correspondence be directed to:

Stephen B. Maebius
FOLEY & LARDNER
Customer Number: 22428
22428
PATENT TRADEMARK OFFICE

Telephone: (202) 672-5569
Facsimile: (202) 672-5399

I UNDERSTAND AND AGREE THAT the foregoing attorneys and agents appointed by me to prosecute this application do not personally represent me or my legal interests, but instead represent the interests of the legal owner(s) of the invention described in this application.

I FURTHER DECLARE THAT all statements made herein of my own knowledge are true, and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Name of first inventor	PITIOT Gilles
Residence	151 rue du Chevaleret, 75013 Paris / France
Citizenship	French
Post Office Address	The same as residence
Inventor's signature	
Date	June 13, 2002
Name of second inventor	LACAZETTE Eric
Residence	Appartement 221, 66 boulevard La Fayette, 63000 Clermond-Ferraud / France
Citizenship	French
Post Office Address	The same as residence
Inventor's signature	
Date	June 13, 2002

Name of third inventor	3rd GACHON Francoise
Residence	11 rue des Paillards, 63540 Romagnat / France
Citizenship	French
Post Office Address	The same as residence
Inventor's signature	
Date	June 13, 2002
Name of fourth inventor	
Residence	
Citizenship	
Post Office Address	
Inventor's signature	
Date	

Attny's Docket No. _____

SMALL ENTITY DECLARATION
[37 CFR 1.9(c-f)]

Each undersigned declares that:

(1) ☐ the application attached hereto
(2) ☐ U.S. Application Serial No. _____ filed _____
(3) ☐ U.S. Patent No. _____ Issued _____

is entitled to the benefits of "small entity" status for paying reduced fees under 35 USC 41(a) and (b) to the Patent and Trademark Office by virtue of the following:

(4) ☒ Each undersigned declares that he/she qualifies as an independent inventor, or would qualify had he/she made the invention, as defined in 37 CFR 1.9(c).

(5) ☐ The undersigned declares that he/she is an official empowered to act on behalf of the concern identified below; that this concern qualifies as a small business concern as defined in 37 CFR 1.9(d).

(6) ☐ The undersigned declares the he/she is an official empowered to act on behalf of the organization identified below; that this organization qualifies as a nonprofit organization as defined in

(a) ☐ 37 CFR 1.9(e)(1)
(b) ☐ 37 CFR 1.9(e)(2)
(c) ☐ 37 CFR 1.9(e)(3)
(d) ☐ 37 CFR 1.9(e)(4) State law of France

(7) Each person, concern or organization to which I/we have assigned, granted, conveyed or licensed, or am under an obligation under contract or law to assign, grant, convey, or license any rights in the invention is listed below:

(a) ☐ no such person, concern or organization
(b) ☐ persons, concerns or organizations listed below

[a separate declaration is required from *each* named person, concern or organization having rights to this invention averring to their status as "small entities."]

Full Name PITIOT Gilles

Address 151 rue du Chevaleret, 75013 Paris / France

☒ Individual ☐ Small Business Concern ☐ Nonprofit Organization

I/we acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement of small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I/we hereby declare all statements made herein of his/her own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application, any patent issued thereon, or any patent to which this declaration is directed.

(8) PITIOT Gilles _____ June 13, 2002
Typed Name of Inventor Signature Date

_____ _____ _____
Typed Name of Inventor Signature Date

(9) _____
Name of Small Business Concern or Nonprofit Organization

_____ By _____ _____
Typed Name Signature Date

_____
Title of Signatory

SEQUENCE LISTING

<110> UNIVERSITE D'AUVERGNE CLERMONT I

<120> HUMAN ODORANT-BINDING PROTEINS WHICH BIND HYDROPHOBIC LIGANDS: POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAID POLYPEPTIDES, AND USES THEREOF

<130> D15947

<160> 16

<170> PatentIn Vers. 2.0

<210> 1
<211> 676
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(552)
<223> cDNA396 (676) /gl (hOBPIIa-alpha)

<400> 1

```
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg     54
                                                 Met Lys Thr Leu
                                                   1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg    102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
  5                  10                  15                  20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc    150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
                 25                  30                  35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg    198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
             40                  45                  50

aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc    246
Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
         55                  60                  65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg    294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
     70                  75                  80

gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac    342
Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr
 85                  90                  95                 100

ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac    390
Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
                105                 110                 115
```

```
cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg ggt agg aat   438
Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Gly Arg Asn
            120                 125                 130

cct aat acc aac ctg gag gcc ctg gaa gaa ttt aag aaa ttg gtg cag   486
Pro Asn Thr Asn Leu Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln
        135                 140                 145

cac aag gga ctc tcg gag gag gac att ttc atg ccc ctg cag acg gga   534
His Lys Gly Leu Ser Glu Glu Asp Ile Phe Met Pro Leu Gln Thr Gly
    150                 155                 160

agc tgc gtt ctc gaa cac taggcagccc ccgggtctgc acctccagag          582
Ser Cys Val Leu Glu His
165                 170

cccaccctac caccagacac agagcccgga ccacctggac ctaccctcca gccatgaccc 642

ttccctgctc ccacccacct gactccaaat aaag                             676
```

<210> 2
<211> 170
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
  1               5                  10                  15

Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
            20                  25                  30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
        35                  40                  45

Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
    50                  55                  60

Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
65                  70                  75                  80

Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
                85                  90                  95

Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
            100                 105                 110

Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
        115                 120                 125

Val Gly Arg Asn Pro Asn Thr Asn Leu Glu Ala Leu Glu Glu Phe Lys
    130                 135                 140

Lys Leu Val Gln His Lys Gly Leu Ser Glu Glu Asp Ile Phe Met Pro
145                 150                 155                 160

Leu Gln Thr Gly Ser Cys Val Leu Glu His
```

165 170

<210> 3
<211> 725
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(480)
<223> cDNA396 (725) /SM12 (hOBPIIa-beta)

<400> 3
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
Met Lys Thr Leu
1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg 102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
5 10 15 20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc 150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
25 30 35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg 198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
40 45 50

aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc 246
Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
55 60 65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg 294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
70 75 80

gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac 342
Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr
85 90 95 100

ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac 390
Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
105 110 115

cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg ggg ccg tgc 438
Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Gly Pro Cys
120 125 130

cgc tgt ccc cac gtc ggc tca cct ggc cac ctc acc tgc agg 480
Arg Cys Pro His Val Gly Ser Pro Gly His Leu Thr Cys Arg
135 140 145

taggaatcct aataccaacc tggaggccct ggaagaattt aagaaattgg tgcagcacaa 540

gggactctcg gaggaggaca ttttcatgcc cctgcagacg ggaagctgcg ttctcgaaca 600

ctaggcagcc cccgggtctg cacctccaga gcccacccta ccaccagaca cagagcccgg 660

accacctgga cctacccctcc agccatgacc cttccctgct cccacccacc tgactccaaa 720

taaag 725

<210> 4
<211> 146
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
  1               5                  10                  15

Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
             20                  25                  30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
         35                  40                  45

Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
     50                  55                  60

Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65                  70                  75                  80

Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
                 85                  90                  95

Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
            100                 105                 110

Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
        115                 120                 125

Val Gly Pro Cys Arg Cys Pro His Val Gly Ser Pro Gly His Leu Thr
    130                 135                 140

Cys Arg
145
```

<210> 5
<211> 741
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(726)
<223> cDNA396 (741) /SM4 (hOBPIIa-gamma)

<400> 5

```
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg    54
                                                   Met Lys Thr Leu
                                                     1
```

```
ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg   102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
  5                  10                  15                  20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc   150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
                 25                  30                  35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg   198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
             40                  45                  50

aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc   246
Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
         55                  60                  65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg   294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
     70                  75                  80

gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac   342
Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr
 85                  90                  95                 100

ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac   390
Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
                105                 110                 115

cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg gca tct gct   438
Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Ala Ser Ala
            120                 125                 130

ccc tgc agg gcc gtg ccg ctg tcc cca cgt cgg ctc acc tgg cca cct   486
Pro Cys Arg Ala Val Pro Leu Ser Pro Arg Arg Leu Thr Trp Pro Pro
        135                 140                 145

cac ctg cag gta gga atc cta ata cca acc tgg agg ccc tgg aag aat   534
His Leu Gln Val Gly Ile Leu Ile Pro Thr Trp Arg Pro Trp Lys Asn
    150                 155                 160

tta aga aat tgg tgc agc aca agg gac tct cgg agg agg aca ttt tca   582
Leu Arg Asn Trp Cys Ser Thr Arg Asp Ser Arg Arg Arg Thr Phe Ser
165                 170                 175                 180

tgc ccc tgc aga cgg gaa gct gcg ttc tcg aac act agg cag ccc ccg   630
Cys Pro Cys Arg Arg Glu Ala Ala Phe Ser Asn Thr Arg Gln Pro Pro
                185                 190                 195

ggt ctg cac ctc cag agc cca ccc tac cac cag aca cag agc ccg gac   678
Gly Leu His Leu Gln Ser Pro Pro Tyr His Gln Thr Gln Ser Pro Asp
            200                 205                 210

cac ctg gac cta ccc tcc agc cat gac cct tcc ctg ctc cca ccc acc   726
His Leu Asp Leu Pro Ser Ser His Asp Pro Ser Leu Leu Pro Pro Thr
        215                 220                 225

tgactccaaa taaag                                                  741
```

<210> 6
<211> 228
<212> PRT
<213> Homo sapiens

<400> 6
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
1 5 10 15

Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
20 25 30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
35 40 45

Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Asn Leu Glu Ala
50 55 60

Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
65 70 75 80

Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
85 90 95

Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
100 105 110

Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
115 120 125

Val Ala Ser Ala Pro Cys Arg Ala Val Pro Leu Ser Pro Arg Arg Leu
130 135 140

Thr Trp Pro Pro His Leu Gln Val Gly Ile Leu Ile Pro Thr Trp Arg
145 150 155 160

Pro Trp Lys Asn Leu Arg Asn Trp Cys Ser Thr Arg Asp Ser Arg Arg
165 170 175

Arg Thr Phe Ser Cys Pro Cys Arg Arg Glu Ala Ala Phe Ser Asn Thr
180 185 190

Arg Gln Pro Pro Gly Leu His Leu Gln Ser Pro Pro Tyr His Gln Thr
195 200 205

Gln Ser Pro Asp His Leu Asp Leu Pro Ser Ser His Asp Pro Ser Leu
210 215 220

Leu Pro Pro Thr
225

<210> 7
<211> 607
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(483)
<223> cDNA396 (607) - short form (hOBPIIa-delta)

<400> 7
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
Met Lys Thr Leu
1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg 102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
5 10 15 20

gag gag gag gat gag gga gaa tcg gtg cat cca gaa gaa aat cct gat 150
Glu Glu Glu Asp Glu Gly Glu Ser Val His Pro Glu Glu Asn Pro Asp
25 30 35

gcg gaa gac gga gga gcc tgg caa att cag cgc cta tgg ggg cag gaa 198
Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu Trp Gly Gln Glu
40 45 50

gct cat ata cct gca gga gct gcc cgg gac gga cga cta cgt ctt tta 246
Ala His Ile Pro Ala Gly Ala Ala Arg Asp Gly Arg Leu Arg Leu Leu
55 60 65

ctg caa aga cca gcg ccg tgg ggg cct gcg cta cat ggg aaa gct tgt 294
Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Leu His Gly Lys Ala Cys
70 75 80

ggc atc tgc tcc ctg cag ggc cgt gcc gct gtc ccc acc ttg gct cac 342
Gly Ile Cys Ser Leu Gln Gly Arg Ala Ala Val Pro Thr Leu Ala His
85 90 95 100

ctg gcc acc tca cct gca ggt agg aat cct aat acc aac ctg gag gcc 390
Leu Ala Thr Ser Pro Ala Gly Arg Asn Pro Asn Thr Asn Leu Glu Ala
105 110 115

ctg gaa gaa ttt aag aaa ttg gtg cag cgc aag gga ctc tcg gag gag 438
Leu Glu Glu Phe Lys Lys Leu Val Gln Arg Lys Gly Leu Ser Glu Glu
120 125 130

gac att ttc atg ccc ctg cag acg gga agc tgc gtt ctc gaa cac 483
Asp Ile Phe Met Pro Leu Gln Thr Gly Ser Cys Val Leu Glu His
135 140 145

taggcagccc ccgggtctgc acctccagag cccaccctac caccagacac agagcccgga 543

ccacctggac ctaccctcca gccatgaccc ttccctgctc ccacccacct gactccaaat 603

aaag 607

<210> 8
<211> 147
<212> PRT
<213> Homo sapiens

<400> 8
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
1 5 10 15

Ser Phe Thr Leu Glu Glu Glu Asp Glu Gly Glu Ser Val His Pro Glu
20 25 30

Glu Asn Pro Asp Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu
35 40 45

Trp Gly Gln Glu Ala His Ile Pro Ala Gly Ala Ala Arg Asp Gly Arg
50 55 60

Leu Arg Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Leu His
65 70 75 80

Gly Lys Ala Cys Gly Ile Cys Ser Leu Gln Gly Arg Ala Ala Val Pro
85 90 95

Thr Leu Ala His Leu Ala Thr Ser Pro Ala Gly Arg Asn Pro Asn Thr
100 105 110

Asn Leu Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln Arg Lys Gly
115 120 125

Leu Ser Glu Glu Asp Ile Phe Met Pro Leu Gln Thr Gly Ser Cys Val
130 135 140

Leu Glu His
145

<210> 9
<211> 676
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(552)
<223> cDNA2098 (676) - conventional form (hOBPIIb-alpha)

<400> 9
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg 54
Met Lys Thr Leu
1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg 102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
5 10 15 20

gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc 150
Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
25 30 35

gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg 198
Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val

```
             40                  45                  50

aag gtg aca gcc ctg ggc ggt ggg aag ttg gaa gcc acg ttc acc ttc   246
Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr Phe Thr Phe
         55                  60                  65

atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg   294
Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
     70                  75                  80

gag gag cct ggc aaa tac agc gcc tat ggg ggc agg aag ctc atg tac   342
Glu Glu Pro Gly Lys Tyr Ser Ala Tyr Gly Gly Arg Lys Leu Met Tyr
 85                  90                  95                 100

ctg cag gag ctg ccc agg agg gac cac tac atc ttt tac tgc aaa gac   390
Leu Gln Glu Leu Pro Arg Arg Asp His Tyr Ile Phe Tyr Cys Lys Asp
                105                 110                 115

cag cac cat ggg ggc ctg ctc cac atg gga aag ctt gtg ggt agg aat   438
Gln His His Gly Gly Leu Leu His Met Gly Lys Leu Val Gly Arg Asn
            120                 125                 130

tct gat acc aac cgg gag gcc ctg gaa gaa ttt aag aaa ttg gtg cag   486
Ser Asp Thr Asn Arg Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln
        135                 140                 145

cgc aag gga ctc tcg gag gag gac att ttc acg ccc ctg cag acg gga   534
Arg Lys Gly Leu Ser Glu Glu Asp Ile Phe Thr Pro Leu Gln Thr Gly
    150                 155                 160

agc tgc gtt ccc gaa cac taggcagccc ccgggtctgc acctccagag          582
Ser Cys Val Pro Glu His
165                 170

cccaccctac caccagacac agagcccgga ccacctggac ctaccctcca gccatgaccc 642

ttccctgctc ccacccacct gactccaaat aaag                             676


<210> 10
<211> 170
<212> PRT
<213> Homo sapiens

<400> 10
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
  1               5                  10                  15

Ser Phe Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
             20                  25                  30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
         35                  40                  45

Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala
     50                  55                  60

Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
```

```
 65                  70                  75                  80

Met Arg Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Tyr Gly Gly Arg
                85                  90                  95

Lys Leu Met Tyr Leu Gln Glu Leu Pro Arg Arg Asp His Tyr Ile Phe
            100                 105                 110

Tyr Cys Lys Asp Gln His His Gly Gly Leu Leu His Met Gly Lys Leu
        115                 120                 125

Val Gly Arg Asn Ser Asp Thr Asn Arg Glu Ala Leu Glu Glu Phe Lys
    130                 135                 140

Lys Leu Val Gln Arg Lys Gly Leu Ser Glu Glu Asp Ile Phe Thr Pro
145                 150                 155                 160

Leu Gln Thr Gly Ser Cys Val Pro Glu His
                165                 170
```

<210> 11
<211> 782
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(537)
<223> cDNA2098 (782) - long form (hOBPIIb-beta)

<400> 11

```
cgc cca gtg acc tgc cga ggt cgg cag cac aga gct ctg gag atg aag   48
Arg Pro Val Thr Cys Arg Gly Arg Gln His Arg Ala Leu Glu Met Lys
  1               5                  10                  15

acc ctg ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc   96
Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe
            20                  25                  30

acc ctg gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg   144
Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met
        35                  40                  45

gtg gtc gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc   192
Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser
    50                  55                  60

cca gtg aag gtg aca gcc ctg ggc ggt ggg aag ttg gaa gcc acg ttc   240
Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr Phe
65                  70                  75                  80

acc ttc atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg   288
Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg
                85                  90                  95

aag acg gag gag cct ggc aaa tac agc gcc tgc ttg tcc gca gtc gag   336
Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Cys Leu Ser Ala Val Glu
```

```
            100                 105                 110

atg gac cag atc acg cct gcc ctc tgg gag gcc cta gcc att gac aca   384
Met Asp Gln Ile Thr Pro Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr
        115                 120                 125

ttg agg aag ctg agg att ggg aca agg agg cca agg att aga tgg ggg   432
Leu Arg Lys Leu Arg Ile Gly Thr Arg Arg Pro Arg Ile Arg Trp Gly
    130                 135                 140

cag gaa gct cat gta cct gca gga gct gcc cag gag gga cca cta cat   480
Gln Glu Ala His Val Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His
145                 150                 155                 160

ctt tta ctg caa aga cca gca cca tgg ggg cct gct cca cat ggg aaa   528
Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys
                165                 170                 175

gct tgt ggg taggaattct gataccaacc gggaggccct ggaagaattt           577
Ala Cys Gly

aagaaattgg tgcagcgcaa gggactctcg gaggaggaca ttttcacgcc cctgcagacg 637

ggaagctgcg ttcccgaaca ctaggcagcc ccegggtctg cacctccaga gcccacccta 697

ccaccagaca cagagccccgg accacctgga cctaccctcc agccatgacc cttccctgct 757

cccacccacc tgactccaaa taaag                                       782


<210> 12
<211> 179
<212> PRT
<213> Homo sapiens

<400> 12
Arg Pro Val Thr Cys Arg Gly Arg Gln His Arg Ala Leu Glu Met Lys
  1               5                  10                  15

Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe
             20                  25                  30

Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met
         35                  40                  45

Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser
     50                  55                  60

Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr Phe
 65                  70                  75                  80

Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg
                 85                  90                  95

Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Cys Leu Ser Ala Val Glu
            100                 105                 110

Met Asp Gln Ile Thr Pro Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr
```

```
        115                 120                 125

Leu Arg Lys Leu Arg Ile Gly Thr Arg Arg Pro Arg Ile Arg Trp Gly
    130                 135                 140

Gln Glu Ala His Val Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His
145                 150                 155                 160

Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys
                165                 170                 175

Ala Cys Gly
```

<210> 13
<211> 542
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (43)..(297)
<223> cDNA2098 (542) - short form (OBPIIb-gamma)

<400> 13

```
cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg    54
                                                  Met Lys Thr Leu
                                                    1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg   102
Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
  5                  10                  15                  20

gag gag gag gat gag gga gga tcg gtg cat cca gaa gaa aat cct gat   150
Glu Glu Glu Asp Glu Gly Gly Ser Val His Pro Glu Glu Asn Pro Asp
                25                  30                  35

gcg gaa gac gga gga gcc tgg caa att cag cgc cta tgg ggg cag gaa   198
Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu Trp Gly Gln Glu
            40                  45                  50

gct cat ata cct gca gga gct gcc cag gag gga cca cta cat ctt tta   246
Ala His Ile Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His Leu Leu
        55                  60                  65

ctg caa aga cca gca cca tgg ggg cct gct cca cat ggg aaa gct tgt   294
Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys Ala Cys
    70                  75                  80

ggg taggaattct gataccaacc gggaggccct ggaagaattt aagaaattgg        347
Gly
 85

tgcagcgcaa gggactctcg gaggaggaca ttttcacgcc cctgcagacg ggaagctgcg 407

ttcccgaaca ctaggcagcc cccgggtctg cacctccaga gcccacccta ccaccagaca 467
```

cagagcccgg accacctgga cctaccctcc agccatgacc cttccctgct cccacccacc 527

tgactccaaa taaag 542

<210> 14
<211> 85
<212> PRT
<213> Homo sapiens

<400> 14
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
1 5 10 15

Ser Phe Thr Leu Glu Glu Glu Asp Glu Gly Gly Ser Val His Pro Glu
20 25 30

Glu Asn Pro Asp Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu
35 40 45

Trp Gly Gln Glu Ala His Ile Pro Ala Gly Ala Ala Gln Glu Gly Pro
50 55 60

Leu His Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His
65 70 75 80

Gly Lys Ala Cys Gly
85

<210> 15
<211> 10664
<212> DNA
<213> Homo sapiens

<220>
<223> hOBPIIa Gene

<400> 15
ctctgcttct atattcggcc tccaaatctg tgctcaatgc agcaactgga gtgacccctt 60

aaatacgtaa gtcacagctt gcctttgtca gagctatcca gggtctttca ctcagagcag 120

aagctgaagt cctcgtggtg gtccttaatc cctacatggc cgttccaccc actccccagc 180

ctcatgtgtg gccggtctcc ctggatcatt tgctgtggct gctctgccgt gtgttcccgg 240

aaactgccag cgttctccca cctctgggct ggcactggat gctcctgcca cctggactcc 300

tcttccagct gacgagctca tggcttgctt ccttcatgtc ttaaattcgg tgtttgaatg 360

ccaccttggc gaggctgttc ctcatcaatt catgtaaagg acaaataacc ccttgtctgg 420

ccactcctgt ccaacttgtc cactttgctt tttccatagc actgatcacc atttaaaata 480

atgtatgaac cagtgaccca tcagtgacca gagacagtag attgaacagg tattgcttgc 540

acagtgaaca atcagctcta gaggtttaaa gagaaccaca aagaatggct caggattgca 600

cgagggcagt taaggaagaa ataaacaggg gtggggatcc tttccacaag tgtgggttca 660
gaccgcatgg gagaaggtgt ggttctccca agggaagctg gagaagtttg ctgggttgtc 720
ccagccacag ctggcccacg gtcagggcag aggccagcag gcagggaaat gtaggctggg 780
tctggcaagc aggaagcctc tcttcccccg caaggcaggt gggctggggc tgggaagcct 840
acaagagtca ctgggaaccc acaggtgcag atgccactga atctcaatag gaagccatct 900
ggggggggtcc cctgaattca atgtggtgtg tgccgccaga cgtccccaac ttgtgccact 960
gccatttata caggaagaga aagaaaaagg aagaaatgga agcatctgga accagtcatc 1020
ctggacccat gctaggaggt gcttcccccct acgcctcaac caacaaagct tagcattgca 1080
ccaggtgcac aagagaagcg cctgctggct ccagctccat tgcctcggaa ccagccatga 1140
agggtgcgtg tggagctgga ggcaagacat tgatagctgg cactgcaatt cacttattta 1200
ttttgttcat tttaagtccc ctgcacctag aatataagcc ccccaagcac aggacatttg 1260
ttttattgat cgatgtattc cttgtgcccc aaagaatgag aggcatctag aaagtctgca 1320
aaaatcaaac ataaaaatga acctttattc agtcattgtt attttgatgg atatttgagg 1380
catttccaag atttgtaagt aacaattgaa cccttctcct ggttcatgtg tgggagtgta 1440
tctgttcaaa tgatgcttct gagtggagtt gctgagtctt tggctctagg ttttttttt 1500
ttaagcattt atgctcattg tggtttttaa attaaacatt taaccctgag acactgtaga 1560
ttcccatgca attgtaaaaa gccacacaga gctatcatgt gtatcttcac ctggcttgct 1620
ccagccccaa ccccagcaat gacagacctg ttctccactc ctgcaatctg ctcattgcaa 1680
gaatgtcgtc tattgcaatc ataaaattgt gggattggct tttttttcc tgtgcagcat 1740
cattctctgg agattcatcc tattgttgca tttatcaata gtttattcca ttttacttct 1800
gagtagtgct ctatggtatg gatgtaccac agtctgttta accattcacc tgttggagga 1860
agtctgtgtt tatagatttg ggctatgaca catgtatagg tttttgcatg gacatcagtt 1920
ttcatttccc tgggacaaag gcccaggggt tctattgctg gattctatgc ttgttacagg 1980
gctcattttg ttttgttttg ctttgtttta ttttgttttt aacctgtcaa gccatttttcc 2040
agatccagtt tccatgcatc ctcaccaggc ttcagtatga tcactatgat cttatctcag 2100
ccaccttaat aggtatgtac tgatatatca tggcttttat ttgcatttca ctgatgacta 2160
atggtgttga gcatcttttc atgtgtttat ttgccatctg tatatcctct ctagtcaagt 2220
gtcccttcat gtcttttgtt tacgttctat ttttgaaact gttgagtttt gaaaaattct 2280
ttataaagta tagaaactaa ttctttgttg aatatgtagt ttgtcaatat tttctttcag 2340

tctttggctt gtctttttat tctgttaaca gggtctctta cagagcaaaa ggtttttatt 2400
ttgatgaagt ctattttaac aatttttcct tttatggatc atatttttgg caacaaatct 2460
aaaacctcct gacccagctc catatgtcaa agattttctt gttttctaaa agttttatag 2520
ttttaagttt tatgttgaag tctatgatcc attttgagtt aattttcata cagggtgtga 2580
gaccaaggtt gtggttcttc tttttctttg gtggttttgt ctgtggatgt ccagttgctc 2640
cagcccattt gctacaaaag ctatctttcc tccactgaat tacttttgca tctttgtaaa 2700
aatgtaattg ggtgtatttg tacaggtctg tttgaggatt ctttattgtc tgtcccattg 2760
atctatgcat ctgtccatgt gctagctata taagtcttga aagggtagcc tgtagctggg 2820
tgtggtggcg agagcctgta atcccagcta ttcgggaagt ggaggcagga gaatcgattg 2880
aacccaggag gtggaggttg cagtgagtca agatcgtgcc actgcactcc agcctgggtg 2940
acagagtgaa ctccatataa aaaaataaaa acaaaaataa agtagcgtga tttctcccac 3000
cttattcttt tttaaaaaag ttttagctat tccagttcct ttgcctttcc atataaattt 3060
tagaataatc ttgtctatat ctaccacaat tctttctgga attttgatag aaattttgtt 3120
acatctttat attatttgag agaaatgata tttttactat gttgagtctt ctaatccatg 3180
gatataacgt ctctccattt gtttagatct tctttgattt attttataat cattgcattg 3240
ttttcagcat acaaatccag catatgggtt gttagactta tgcctaggca tctcattttt 3300
ttagccatta taaatagtac tgtgttttta agtttagggt ccattactag tatgtagaca 3360
cacaattgat ctttgtatat ttatcttgtg tcctgcaacc ttgctgaact cacttactag 3420
ttctaggagg tgtattgttt tcttttgttt tgtttttcaa tttcttggga ttttctacag 3480
agataataat gtcatctgca gatgcagttt tcttccttcc tttccaattt gtatgtcttt 3540
aatttccttt ttaaaaaacc tatattactc tgactagaac tttctgtact atgttaaata 3600
caagtggtag agtggacatc cttgccttgt ccctgatgtt aaagagaaag catttgtaac 3660
tgagtatccc agcctcaaaa tgtgcttaaa aacttttttc ctttcttgct ttcagccttg 3720
aaacatactt cgaaactctt tatttctccc tttcccacca ggcacttccg tgagcagtgc 3780
tcgcttatct aattatgtgc ttacttagaa attccagggg ccaattttga aacaaaccag 3840
gcagagagac ccagctgcag aatcctccct cttaggggga gttacaggta gcctaccact 3900
tcccggctga aatcaggatg acgcaaacca gacctccgga cagacgattg atgactcaca 3960
ataaccatca gaacaagatg cagaccaaca tcctcctgca ccattcccac atatttccca 4020
caccttttcc tccttaaacc ccttcgctca gtccagaaaa tctgaatggt cttttaaagg 4080

catgggtctg gccattcccc aactgccagt atttgaataa agctgctttc cttttaccac 4140
acctcacttc tcatgccttg acttctgagc agcgagcagc tggacttgag ccagttacac 4200
atcagtcttt caccattaga aataatgtag ccataggttt ttttttcgta gatgttcttt 4260
ttcaagttaa agaagttctc ttctattcct atttttctga gaggtgttat cccgaatgag 4320
tgttgaattt tgttaaatac ttttaacaac ccaacaggaa ccaccatcag aagccatcca 4380
gagaaacgca ccaggccaca aagcactggg ggccagggat cttgcccctg ctgtctgcca 4440
tgggttgacc cccagcctcc aaccctacca tcccctgacg gtgtctgcag cagttgaacc 4500
caaccagcat ctaaaagaac acagttggtg aacgagactg ggacacaggg caagatgggt 4560
ggacaatggg aggctcctgg agagcacccg taccagcgag catagaattc atgggggtga 4620
cctgttccct gaagcatctg cgcgtgttgt tccagcattt tcttcaagga ttgagccagc 4680
agcaccagtg tcatacggtg cttaaatcaa tgattcacag ccaaccaatg aaatacaagg 4740
tgccggctgg gcgcggtggc tcacgcctgt aatcccagca ctttgggagg ccgaggcggg 4800
cggatcacga ggtcaggaga tcgagaccat cccggctaaa acggtgaaac cccgtctcta 4860
ctaaaaatac aaaaaaaatt agccgggcgt agtggcgggc gcctgtagtc ccagctactt 4920
gggaggctga ggcaggagaa tggcgtgaac ccgggaggcg gagcttgcag tgagccaaga 4980
tcccgccact gcactccagc ctgggcgaca gagcgagact ccgtctcaaa aaaaaaaaaa 5040
aaaaaaaaaa aaaagaaata caaggtgcct ggggtacagg caacaaaaat ggggaacgga 5100
gaaatcttcc cgttgaggta gagactccag cctgtttgct catctctggt cctgaagagc 5160
ccagtccccg ccctaaggat ggggtttctg ctcggcagca cttgccgtga gaggggtgag 5220
gcactgggtc acgccagccc tttctttata gccccccaggg tttattagag tgtgcattag 5280
tattatttag caagcacttg agggtgtctg atcttgggcc agagaggcac aaagattgtg 5340
tgggcccagc acctgcctgc agagcgtggg tcagccttgg gtcccagggc agatgacgca 5400
ggcccgggaa ctacagggtc caggagggag aaggcaaagt tctggctcag gttggctggg 5460
gatgaggcca gcggagccag gtgcccaagg gagctcagcc acaaactctg agcacaggct 5520
ggcaggtggc tcttgatgct catgccaccc atttatctaa agggatgaga ttcaaggcct 5580
gtcctggtgg ctgggggccg tcgaagctga cgagagaggg gatgtagagt gaatgtatat 5640
tccactctac cacttgtatt taacagggag atggatgatg aacacgtgca ggaggaaaca 5700
ggcaggacaa tccagagaga tcacgtgttc tgaggacagc acagccaggc tccggtacgg 5760
agtgaagcgg ggtggggcag gcggcggggt ccctcatatg gcccgaggag gccgtatata 5820

tactgacctt gagccacaca atagtgccct tctctgcccc taggaagctc gagtgcaggg 5880
tccaggtggg gaaaatcaat gcagagtggg tcccagagtg ggcggaagct tgggctctag 5940
ggcgtgcggg actcagctgg cagcagcccc acatctctat ggttctaaag cccagtccat 6000
ttctgctcag cagggagatg gccagttccc cagaggactt gcccagggtc ccagccgtgg 6060
ccctgggagg gctcaggggt gagggaggga gaattccgag ccgtgggccc tccctctgtg 6120
gctcggggtg cagacgtgca ctaccctgcc ctgtcctctg gagtcctccg ggctcatcca 6180
agtgggcgca tttgggtgca aaccctggac agcatcgtgt gtcttttctt ctctggctgg 6240
cactgaattg ctgttcacca gatgccagca atgaggacac tccccctccc ggatgcagga 6300
ctggttgcca ccatggggag gggtgcaccg tgccacgtgc tcccaggaca ctggccaggg 6360
ggctataaag aacatctcga gaggagccag cacagccttg ttcagacgcc cagtgacctg 6420
ccgaggtcgg cagcacagag ctctggagat gaagaccctg ttcctgggtg tcacgctcgg 6480
cctggccgct gccctgtcct tcaccctgga ggaggaggat gtgagctggg ttggcgtggg 6540
cggatggagg agccaggtgg actcctgggc aggggcagt gccaggggcc ctgctttagg 6600
aggtgtcact taagcctggg gtggtggtgg aggggtccta ttgttccttc agcagacaat 6660
gctccccatg aggcccaggg ccggagcagg ctcggctcag gggttcctgc tgcactgacg 6720
cctgaagccc gaaggtctcg cagggttggg ccctgtggag ggagggctca cctggtgctg 6780
gggcccgggg gtccatgggg tgcagacatg ccctccttcc actgggggct gggagccctg 6840
agcaggggc tggctctaac tcactccagc tgagctctaa ctaaggtgca ggaacccagc 6900
ctgcctttag gggtggcagc cgggcaccat gggtgtctgg ttatagctgc aggcctgagt 6960
gccagggtca gagtagaatc tgggccaccc atggtgggct cacggccttg gcctgctcca 7020
gatcacaggg acctggtacg tgaaggccat ggtggtcgat aaggactttc cggaggacag 7080
gaggcccagg aaggtgtccc cagtgaaggt gacagccctg ggcggtggga acttggaagc 7140
cacgttcacc ttcatgtgag tgttgcccac tgcagggccc ctcaggccac tttcgctccc 7200
cgccccagac ccacctggtg cccattgccc catccacatt tcgggtgttg ggaagagtca 7260
ccccctgcct tggagggaaa cagccagggc atcctgaagc tcggtggggt ggggggcag 7320
tggaattttc aggttgccgg gtcagggcca tgcaccaggt gagctgagga tgggccaggt 7380
gtgtcctggg agccgctgcc cgcgtgtctc ctgttttcca ggagggagga tcggtgcatc 7440
cagaagaaaa tcctgatgcg gaagacggag gagcctggca aatacagcgc ctgtgagccc 7500
ctccccgacc ccccactccc catgcccaac cccggatgca ccagccccac tgcaggtgga 7560

gagtgcccag gccacacttc tgccagggtc ccagccctgc ccacctccaa ggaggggctg 7620

gcctctcctt cctgggggge tggtggccct gacatcagac accaggtgtg acaggcttgt 7680

ccacagtaga gatggaccag atcaagcctg ccctctggga ggccctagcc attgacacat 7740

tgaggaatcc gagtgttagg gaccaggagg ccgagggtta gggatgggaa gccaaggctg 7800

agggtttggg atcaggaatc cgagggttag ggacagggaa gtggggcagg agcagctgct 7860

ggagctggga aggccggact ctagtcctgg acgtgctctg gccttgtggc tccattactt 7920

gcattgggac cttccgagag gaggctcctg cctccgtgtc cgggtccatg ctgtgcggag 7980

cagccaggcc tggctcaggc tgtccagggc acctgggtga ccactgaaac attcctgagt 8040

gtttcttcgt gtggtcctga gtgctctctc cgggaatgag ggcactgaag acccatcttc 8100

tctgtcatct acagatgggg gcaggaagct catatacctg caggagctgc ccgggacgga 8160

cgactacgtc ttttactgca aagaccagcg ccgtgggggc ctgcgctaca tgggaaagct 8220

tgtgggtgag gggcccgctg gggcctgcat gtcctgcccc atggtctctg cctccagaag 8280

ccagtggaac caccatcatc acgccctggc acgggggggaa aaggaagccc cctgcgccgg 8340

ccttcgtgtg ctaggcacca agcgctgccc tggatggctg gtccaagttc ctgaagtggg 8400

agtggggtgg gccaggcagg gacagacacg gccctcggtg acgtgaacct gccaagggcc 8460

gcttgtgggg tctcaggtgt aggggcctca ccttaagggg gaggtancat cttaacagag 8520

ctcttcatgg ggcagggact ctccagggcg gcagggcagc cagtgcctct gggacacaag 8580

gtccctccag gtgagggttg tgaccctgca gagtggcttt gggagctgcc caggtccccc 8640

tggggttgct gagtggcttg gaccctgcca ctgtcccctt tcctggggac ctctcacctg 8700

ggcggtggcc gtctcctctg tccccagtcc caccctgag ctcttgtcca ttctcaggcc 8760

tcctctcccc cttgcctggt gctggacagt tgccatctct tctgtcccca gccccacccc 8820

tgagctctga tccactctcg ggcctctccc ccgtcctgat gctgggccgt ggtcgtctcc 8880

ttttagcatc tgctccctgc agggccgtgc cgctgtcccc acgtcggctc acctggccac 8940

ctcacctgca ggtaggaatc ctaataccaa cctggaggcc ctggaagaat ttaagaaatt 9000

ggtgcagcac aagggactct cggaggagga cattttcacg cccctgcaga cgggtgagga 9060

cggctgtgcc cagtaccccg tgttcccctg tgtctctgtg tgatctccag tgtcccatga 9120

ccctcgtgtc ctcccatgtc ccccgcattc cccatgtgcc ccgagtctcc tcgcaggggc 9180

tctgggccct gcttagcatc ctcgtcgttg gagggtctgc actctgggct gcgatggggt 9240

ctggggctcc gcgctctggg ctgcgatgcg gtctggggct ccgcgctctg ggctgcgatg 9300

gggtctgggg ctccgcgctc tgggctgcga tggggtctgg ggctccgcgc tctgggctgc 9360

gatggggtct ggggctccgc gctctgggct gcgatggggt ctggggctcc gcgctctggg 9420

ctgcgatggg gtctggggct ccgcgctctg ggctgcgatg gggtctgggg ctccgcgctc 9480

tgggctgcga tggggtctgg ggctccgcgc tctgggctgc gatggggtct ggggctctga 9540

gctctgggct gcgatggggt ctgggccctg gtctagggcg ccttctaatc cctgggtttt 9600

tcttggtctc tgcaggaagc tgcgttctcg aacactaggg tgagtgagcc tttaggaggg 9660

cactggacaa gcccagagtc ctgggttccc ggggttcgag ggtacatctg ctctggccct 9720

tcccatccca cacagccagg gagacccccc cagggtcagg cacgaggttg gcacctcaga 9780

gtctgcccac ccaaaattcc tgggacattc gggaagtcct ttgttttacc attcctgcac 9840

ctgccagccc gagtgagggt cctcctcggc ctttccacag cgaggcctcc ctccggctcc 9900

ctcaggtgtc agctcggccc atcgcccccc gcacctgtcc ccactcggtc tactcccccc 9960

cacccactca ccaaggatgc tcagaggcct gcccagtcat tggaggagct gaggctgctc 10020

tggagccccg aggctgccca gcagtggccg atgtggaagc tgcagagcct ggggagggag 10080

ctctgggcct ggcctctgcc ctaccctgca gcctccctga cctctgctcc tcttcccagc 10140

acagcagccc ccgggtctgc acctccagag cccaccctac caccagacac agagcccgga 10200

ccacctggac ctaccctcca gccatgaccc ttccctgctc ccacccacct gactccaaat 10260

aaagagcttc tcccccagct ctgggcaggc ctatctgtgg ggacggaggg gctcgcaccg 10320

gctccctggg aggcctgctg ggagggggag ccacaaggga agctggggag atgttggacc 10380

tagcaagggc caaaccacaa gaggcacgat gtccaacagg ctgtggggct gcgtgatgtc 10440

ctggaggggc ctccggaaga gccgccctcg gatctcaggc tcagcttggg acgggcaggg 10500

ctctgggcag gagcacttgg gaagccactg ggagggccgg agtggggaca cggcgtcaga 10560

cctgcatcag tgggtccacc acagggctcc cgccccgggg gtctcagtgg gatcctccga 10620

gcgggtccac tttancatcc tgngctcagt ggcatcactg gata 10664

<210> 16
<211> 13591
<212> DNA
<213> Homo sapiens

<220>
<223> hOBPIIb Gene

<400> 16
tctctgcttc tatattcggc ctccaaatct gtgctcaatg cagcaactgg agtgacccct 60

taaatacgta agtcacagct tgcctttgtc agagctctcc agggtctttc actcagagca 120
gaagctgaag tcctcgtggt ggtccttaat ccctacatgg cagttccacc cactccccag 180
cctcatgtgt ggccggtctc cctggatcat ttgctgtggc tgctctgctg tgtgttcccg 240
gaaactgcca gcattctccc acctccaggc tggcactgga tgctcctgcc acctggactc 300
ctcttccagc tgacgagctc atggcttgct tccttcatgt cttaaattcg gtgtttgaat 360
gccaccttgg cggggctgtt cctcatcaat tcatttaaag gacaaataac cccttgtcct 420
gacactcctg tccaacttgt ccactttgct ttttccatag cactgatcac catttaaaat 480
aacgtatgaa ccagtgaccc atcaggatcc agagacaata gattgaacag ggattgcttg 540
cacaatgaac aatcagctct agaggtttaa agagaaccac aaagaatggc tcaggattgc 600
acgagggcag ttaaggaaga aataaacagg ggtggggatc ctttccacaa gtctgggttc 660
agacctcatg ggagaaggtg tggttctccc aagggaagct ggagaagttt gctgggttgt 720
cccagccaca gctggcccac ggtcagggca gaggccagca ggcagggaaa tgtaggctgg 780
gtctggcaag caggaagcct ctctccccac ccaaggcagg tgggctgggg ctgggaagcc 840
tgcaagagtc actgagaacc cacaggtgca gatgccactg aatctcaata ggaagccatc 900
tgggggcggt cccctgaatt caatgtggtg tgtgccgcca gacgtccccg acttgtgcca 960
ctgccatttg cataggaaga gaaagaaaaa ggaagaaatg gaagcatctg gaaccagtca 1020
tcctggaccc atgctaggag gcgcttctcc ctacgcctca accaacaaag cttagcattg 1080
caccaggtgc acaagagaag cgcctgcttg ctccagctcc attgcctcgg aaccggccat 1140
gaagggtacg tgtggagctg gaggcaagac attgatagct ggcactgcaa ttcacttatt 1200
tattgtgttc attttaagtc ccctgcacct agaatataag ccccccgagc acaggacatt 1260
tgtttcattg atcgatgtat tccttgtgcc ccaaagaatg agaggcatct agaaagtctg 1320
caaaaatcaa acataaaaat gaacctttat tcagtcattg ttattttgat ggaaatttga 1380
ggcatttcca agatttgtaa gtaacaattg aacccttctg ctggttcatg tgtgggagtg 1440
tatctgttga aatgatgctt ctgagtggag ttgctgagtc tttggctcta ggttttttt 1500
ttttaagcat ttatgcttat tgtggttttt aaattaaaca tttaaccctg agacattgta 1560
gattcccatg cagttgtaaa aagccacaca gagctatcat gtgtatcttc acctggcttg 1620
ctccagcccc aaccccagca atgacagacc tgttctccac tcctgaaatc tgctcattgc 1680
aagaatgtcg tctattgcaa tcataaaatt gtgggattgg ctttttttt tcctgtgcag 1740
catcattctc tggagattca tcccattgtt gcatttatca atagtttatt ccattttact 1800

tctgagtagt gctctatggt atggatgtac cacagtctgt ttaactattc acctgttgga 1860
ggatgtctgt gtttatagat ttgggctatg acacatgtac aggtttttgc atggacatca 1920
gttttcattt ttctgggaca aaggcccagg ggttctattg cggggttcta tgcttgttgc 1980
agggtttttt tttttaaacc tgtcaagcca ttttccagaa actgtccagt ttccatgcat 2040
cctcaccagg cttcagtatg atcactatga tcttatctca gccaccttaa taggtatgta 2100
ctgatatatc atggctttta tttgcatttc actgatgact aatggtgttg agcatctttt 2160
catgtgttta tttgccatct gtatatcctc tctagtcaag tgtcccttca tgtcttttgt 2220
ttacattcta tttttgaaac tgttgagttt tgaaaaattc tttataaagt atagaaacta 2280
attctttgtt gaatatgtag tttgtcaata ttttctttca gtctttggct tgtcttttta 2340
ttctgttaac agggtctctt acagagcaaa aggtttttat tttgatgaag tctattttaa 2400
caatttttcc ttttatggat catatttttg gtgacacatc taaaatctcc tgacccagct 2460
ccatatccca aagattttct ttctgttttc taaagctttt atagttttaa gttttatgtt 2520
taagtctatg atccattttg agttaatttt cttgcagggt gtgagaccaa ggttgtggtt 2580
cttctttttg tttggtggtt ttgtctgtgg atgtccagtt aatccagccc atttgctaca 2640
aaagctatct tccattgaat tgcttttgca tctttgtaaa aacttaactg ggtatatgcg 2700
tgcaggtctg tttgaggatt ctttattgtc tgtcccattg atctatgcat ctgtccatct 2760
gctagctata taatgagtct tgaaagggta gcctgtagct gggtatggtg gtgtgcatct 2820
gtaatcccag ctacttggga ggcggaggca ggagaatcac ttgaatctgg gaggtggagg 2880
ttgcaatgag tcgagattgt gctactgcac tccagcctgg gtgacagagc aagactctgt 2940
ctctaaaaaa aagaaagaag aagaaaagga aagaaagggt agcctgattc ctcccacttt 3000
attcttaaaa aaaaaagttt tagttattct agttcctttg cctttccata taaatgttag 3060
aataatcttg tctatgtcta caaaaattcc ttctggaatt ttgatagaaa ttgtgttaaa 3120
tctttatatt atttgagaga aatgacgctt ttattatgtt gagtctccta atccatggat 3180
atagcacatc tctccatttg tttagatctt ctttgattta ttttataatc attgcattat 3240
tttcagcata caaatccagc atatgggttg ttaaacttat gcctaggtat ctctttttt 3300
tagccactat aagtagtatt gtgtttttaa gtttagggtc cgttactagt atgtagacac 3360
acaattgatc tttgtatatt tatcttgtat cctgcaacct tgctgaactc gcttaatagt 3420
tctaggaggt gtattgtttt cttttgtttt gtttttcagt ttcttgggat tttctacaga 3480
gacaatcatg tcatctgcag atgcagacag ttttcttcct tcctttccaa tttgtatgtc 3540

tttaatttcc tttttaaaaa acctttattg ctctgactag aactttctgt actatgttaa 3600
atacaagtgg tgaaagtgga catccttgcc ttgtccctga tgttaaggag aaagcgtttg 3660
taactgagta tcccagcctc aaaatgtgct taaaaacttt tttccttttct tgctttcagc 3720
cttgaaacat acttcgaaac tctttatttc tccctttccc accaggcact tctgtgagca 3780
gtgctcgctt atctaattat gtgcttactt agaaattcca ggggccaatt ttgaaacaaa 3840
ccaggcagag agacccagct gcagaatcct ccctcttagg gggagttaca ggtagcctac 3900
cacttcccgg ctgaaatcag gatgactcaa accagacctc tggacagacg attaatgact 3960
catgataacc attggaacaa gatgcagacc aacatcctcc tgcaccattc ccacatattt 4020
cccacacctt ttcctcctta aacccccttcg ctcagtccag aaagtttgaa tggtctttta 4080
aaggcatggg gcctggccat tcctctactg ctagcatttg aataatgctg ctttcctttc 4140
accacacttc acttctcatg ccttgacttc tgagcagcga gcagctggac ttgagccagt 4200
tacacatcag tctttcacca tgagaaatga tgttagccat aggtttttg tagatgctct 4260
ttgtcaagtt aaggaggttc tcttctattc ctacttttct gagagttgtt ttcctgaatg 4320
cgtgatgaat ttagtcaagt actttttctg cattgatcga tatgatcatg taatttttct 4380
tcttaacata ttaatatggt tgatttttga gtattgtacc aggtttccat ccctggaata 4440
aactgctttt ggtcatggtg tagaattaat tttatatatt attatttgct aatattttgt 4500
aaaggatttc tgtatctata ttcatgaggt atattgttct gcagctttcc tttttgtgtt 4560
gtcaggtttt ggttgggaga tattctctca tcttctgaga ttatgcagag ttggtgttat 4620
ttatttttta aatggttggt agaatttttct aatgaaacca tatggacatg aagaattatt 4680
tttggaagct tttaaaatta aaatgttaac tttaattgtc atagagctat taaagttatc 4740
tatcttatat tggcattgtg ttttttgaga aattggttca tttcacctat gttgtcaaat 4800
atatgtggga aggtctcttc gtagtattcc cttattatcc tttgatgtct gcagggtctg 4860
tagtgatagt ctctgttttg ttccagatgt tggttatttg tgtctttctt tctttttttg 4920
tttgtcagtc tttctagaac tgcctcaatt ttattaattt ttccagagaa ctaacccttg 4980
tttcattgtt tttctctgtt gtttttctgc tttcgatttt ttctgctttc tgcttttgat 5040
ttattaattg ctgctcttat ctttgttatt tctttttcttc ttgctttggg tttattttgc 5100
tcttcttttt ctaggttctt gaagtgggag cttagatagt tgattttgga cttctctttt 5160
ctaatatatg catttaatgc ctctcagcac tgatttagct gtgtctcaca aattttgata 5220
tgtttttgtt tgtttgtttt tgaggtagag tctcgctctg tatcccaggc tggagtgcag 5280

tggcactatc ttggctcact gcaacctcca cctcctgggt tgaagtaatt ctcctgcttc 5340
agccccctaa gtagctggga ttacaggtat acgccaccat gcccagcaat ttatttattt 5400
atttattttt gtatttttag tagagatgag gttttgtcat gctgtccagg ctggtctcga 5460
actcctgacc tcaagtgatc tgcccacgtg gggctccaaa gcactgagat tataggcata 5520
agccactgtg cccagcctgt ttttgttttt tttgagatgg agtcttgctc tgtcacccag 5580
gctggagtgc agtggcatga tagcccactg caacctccgc ctcccgggtt caagtgattc 5640
tcctgcctca gcctcccgat tagctgggat tacaggcgtg caccaccaca ctcagctagt 5700
ttttgcattt ttggtagaga tgggttttcg ccatgttggc caggctggtc tcgaattcct 5760
gatcgcaagt gatccacccg cctgggcctc ccaaagtgct ggaattatag gcacaagcca 5820
tcacgcccag ctgatatgtt gtattttcat tttcatccag cctagtatat ttttaaattt 5880
accttcaggc ttcctctttg acctgtggat tattcacagg tgtgttgttt ccaagcatct 5940
gaagaaagat tttccattat ctttctggca ttgatttata gttagattcc atgtggtcag 6000
agactaccct ctatatcatt taaattttta aaaatttctt gaggctcatt ttctcatcca 6060
ggatatggtc tatcttggta tatattctgg tgcacttgaa aagaatgtgt gtagtgctgt 6120
tgccaggtgg tgtgttctac aaatgtcaat tcgattatgt taattgatgg gtggccgagt 6180
tctttgatgt ccctgctggt tttctgtcta gttgctctga gagaagggta ttgaagtctc 6240
caactataat tgtggctttt taaatttctc ctttcagttc tgttttgctt cacatattgt 6300
gcagctaatg tttggtgcac acacgtttag gattactatg tcttctttgc agaatggccc 6360
acttattatt gtataatttc cctctctgtt aattagtctt tgttctaaag tctattttac 6420
ctgttattga aataatgctt ctgctttctt ttgattaatg tttatatatc ttttttcatg 6480
tttttacttt caacctgtct ctattgttac atttgaagta agtttcttgt agacggcata 6540
tagttgtata atggctttta attcactttg ccaatcactg tcatttaata tttaaaccat 6600
ttacatttaa tataattatt aacatgctag agtttgtatt atttttatag ggttgctata 6660
acaaaatact accaactggg tggcttagaa caaaaattta ttttctcaca ttctggtggc 6720
cagaagtctg agatcaagtg agatttggaa ggtagaaaaa ccaacaaaat atgttaatga 6780
tgttaccatt gtgggcaact gggactcagt ccttctgagg gccatctgag aaatagggtg 6840
gaaccatatt tgattgttcc acttagaagg aagagtccag accatttatc ccatcagtta 6900
ctatccctac agactgagga tgctgattca cttgcacatc tgggtttcat cagtgtcagg 6960
ggaagcacag aagaaaaggt gcaggcactt aaggtgggaa gctgtgaata cgtccgtgca 7020

gggctctagg gcgtgcggga ctcagctggc agcagcccca cgtctctctg gttctaaagc 8820
ccagtccatt tctgctcagc agggagatgg ccagttcccc agaggacttg cccagggtcc 8880
cagccgtggc cctgggaggg tgcaggggtt ggtgagggag tatcccgagg ctgtgggccc 8940
tgcctctgtg gcttggggtg cagatggaca ccaccctgcc ctgtcctctc gagtcctctg 9000
ggctcatctg agtgggcgcg ttcgggtggt gcaaactctg gatggcatcg tgtgtctttt 9060
cttctctggc tggcactgac ttgctgttca ccagatgcca gcaatgagga cactcccct 9120
cccggatgca ggactggttg ccaccatggg gaggggtgca ccgtgccacg tgctcccagg 9180
acactggcca gggggctata aagaacatct cgagaggagc cagcacagcc ttgttcagac 9240
gcccagtgac ctgccgaggt cggcagcaca gagctctgga gatgaagacc ctgttcctgg 9300
gtgtcacgct cggcctggcc gctgccctgt ccttcaccct ggaggaggag gatgtgagct 9360
gggttggcgt gggcggatgg aggagccagg cggactcctg ggcaggggggc agtgccaggg 9420
gccctgcttt aggaggtgtc actgaaggct ggcttctgac cccgtgctcc cagcctgggg 9480
tggtggtgga ggggtcctat tgttcctcca gcagacacgg ccccccgagg cccagggccg 9540
gagcaggctc gtctcagggg ttcctgctgc actgacgcct gaagcccgaa ggtctcgcag 9600
ggttgggccc tgtggaggga gggctcacct ggtgctgggg cccgggggtc catggggtgc 9660
agacatgccc tccttccact gggggctggg agccctgagc agggggctgg ctctaactca 9720
ctccagctga gctctaacta aggtgcagga acacagcctg cctttaaggg cagcagccgg 9780
gcaccatggg tgtctggtta tagctgcagg cctgagtgcc agggtcagag tagaatctgg 9840
gccacccatg gtgggctcac ggccttggcc tgctccagat cacagggacc tggtacgtga 9900
aggccatggt ggtcgataag gactttccgg aggacaggag gcccaggaag gtgtccccag 9960
tgaaggtgac agccctgggc ggtgggaagt tggaagccac gttcaccttc atgtgagtgt 10020
tgcccactgc agggcccctc aggccacttt cgctcccctc cccagaccca cctggtgccc 10080
attgccccat ccacgtttcg ggtgttggga agagtcaccc cctgccttgg agggaaacag 10140
cctgggcatc ctgaagctcg gtggggtggg ggacagtgga attttcaggt tgccgggtca 10200
gggccatgca ccaggtgagc tgaggatggg ccatgggtgt cctgggagcc gctgcccgcg 10260
tgtctcctgt tttccaggag ggaggatcgg tgcatccaga agaaaatcct gatgcggaag 10320
acggaggagc ctggcaaata cagcgcctgt gagcccctcc cccactccca cccccaccct 10380
cccccaccgc caaccccagt gcaccagcct ccacaggtag agagtgccca ggctgccctt 10440
ttgccagggc cccagctctg cccacctcca aggaggggct ggcctctcct tcctgggggg 10500

ctggtggccc tgacatcaga caccgggtgt gacaggcttg tccgcagtcg agatggacca 10560
gatcacgcct gccctctggg aggccctagc cattgacaca ttgaggaagc tgaggattgg 10620
gacaaggagg ccaaggatta ggtacgggga ggctgagggt tagggatggg gaagctgagg 10680
gttaaggatg gggaagctga gggttaagga tggggaggct gagggttagg gacgggggcc 10740
aagggttagg gatcgggaag ctgagggtta gggatgagga ggccgagggt tagggatggg 10800
gaggccgagg gttagtgctg cggaagctga ggattagaga tggggaggct gaggggggaag 10860
atgagggtta gggacaggga agtggggcag gagcagctgc tggagctggg aaggccggac 10920
tctagtcctg gacgtgctct ggccttgtgg ctccattact tgcattggga ccttcctgag 10980
aggaggctcc tgcctccgtg tccgggtcca cactgtgcgg agcagccagg cctggctcag 11040
gctgtccagg gcacctgggt gaccgctgaa acattcccga gtgtttcttc atgtggcccc 11100
gagtgttctc tctgggaatg acccattttc tctgtcatct acagatgggg gcaggaagct 11160
catgtacctg caggagctgc ccaggaggga ccactacatc ttttactgca aagaccagca 11220
ccatgggggc ctgctccaca tgggaaagct tgtgggtgag gggcttgctg gggtctgcgt 11280
gtcctgcccc acggtctctg cctctggaag ccggtggaac cagcaccacc atgccctggc 11340
atgggggaaa aggaagcccc ctgtgccggc tttcatgtgc cgggcaccgc tgggcagccg 11400
gtccaagtac ctgaagtggg aatgggatgg gccaggcagg gacagacatg gccctcagtg 11460
acatgaacct gccaagggcc acttctgggg tctcgggtgt aggagggtca ccttaagggg 11520
gagggtcacc ttaagcgtcc taagagagct cttcatgggg cagggaccct ccagggcagg 11580
agggtggccg gtgcctctgg gagacaaggt ccctccaggt gagggctgtg accctgcagg 11640
gtggcattgg gagctgccca ggtcccctgg ggttgccgag tggcttggag agtccccgcc 11700
actgtccccc ttcctgggga cctctcatct gggcagtgac tgtctcctct gtccccagtc 11760
ccacccctga gctctcatcc attctcaagt ctcctctcct gcttgtccgg tgctgggcta 11820
tggccatctc ttctgtcccc ggccccaccc cagagctctg gtccactctt gggtctctcc 11880
ccttgtcctg gtgctgggca gtggccatct cctctgtccc cagccccacc cccgagctct 11940
gatccactct caggcctctc cccctgtcct ggtgctgggc cgtggtcgtc tcctttcggc 12000
atctgctccc tgcagggccg tgccgctgtc cccacgtcgg ctcacctggc cacctcacct 12060
gcaggtagga attctgatac caaccgggag gccctggaag aatttaagaa attggtgcag 12120
cgcaagggac tctcggagga ggacattttc acgcccctgc agacgggtga ggatggctgt 12180
gcccagtccc ctgtgtccct ctgctgtgtc tgtctgctat ctccagtgtc ccatgacccc 12240

catgtcctcc catgtccccc gcattcccca tgtgccccga gtctcctcgc aggggctccc 12300
gggccctgtt tagcgtcctc ctcattggag gctctgtgct ctgggctgcg atggggtctg 12360
gggctccgcg ctctgggctg cgatggggtc tggggctccg cactctgggc tgcgatgggg 12420
tctggggctc cgcgctctgg gctgcgatgg gctctggggc tctgagctct gggctgtgat 12480
ggggtctggg ccctggtcta gggcgccttc taatccctgg gtttttcttg gtctctgcag 12540
gaagctgcgt tcccgaacac tagggtgagt gagcctttag gagggcactg gacaagccca 12600
gagtcctggg ttcccggggt tcgagggtac atctgctctg gcccctccca tcccacacag 12660
ccagggagac ccccccaggg tcaggcacga ggttggcacc tcagagtctg cccacccaca 12720
gttcctggga cattcgggaa gtcctttgtt ttatcattcc tgcacctgcc agcccgagtg 12780
agggtcctcc tcggcctttc cacagcgagg cctccctccg gctccctcag gtgtcagctc 12840
agcccatcgc cccctgcacc tgtccccacg cggtccactc cccaccatcc acccaccaag 12900
gatgctcaga ggcctgccca gtcattggag gagctgaggt cgctctggag ccccgaggtt 12960
gcccagcagt ggccgatgtg ggggctgcag agcctgggga gggagctctg gccctggcct 13020
ctgccctacc ctgcagcctc cctgacctct gctcctcttc ccagcacagc agcccccggg 13080
tctgcacctc cagagcccac cctaccacca gacacagagc ccggaccacc tggacctacc 13140
ctccagccat gacccttccc tgctcccacc cacctgactc caaataaagt ccttctcccc 13200
cagctctggg caggcctatc tgtggggaca gaggggcttg cacccgctcc ctgggaggtc 13260
tgctgggagg gggagccaca agggaagctg gggagatgtt ggacctagca agggccaaac 13320
cacaagaggc acgatgtcca acaggcctga ggggctgcgt gatgtcctgg aggggcctcc 13380
ggaagagccg ccctcaggtc tcaggttcag cttaggacag gcagggccct gggcaggagc 13440
gtttggaaag ccactgggga ggacaggagc ggggacacgg cgtcagggct gcagcagtgg 13500
ggccaccgca gggctcccgc ctcaggggtc tcagagggat cctccaagct cccccatttt 13560
agcagcctgg gctcagtggc agcactggag a 13591

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division

Application deficiencies found during scanning:

☐ Page(s)_____of_____ were not present for scanning. (Document title)

☐ Page(s)_____of_____were not present for scanning. (Document title)

☑ ***Scanned copy is best available.*** Some of the Drawings are very Dark & Some are faded